QUARTERLY OF APPLIED MATHEMATICS

Vol. VIII October, 1950 No. 3

THE DIRICHLET PROBLEM: BOUNDS AT A POINT FOR THE SOLUTION AND ITS DERIVATIVES*

BY

CLAIR G. MAPLE**

Carnegie Institute of Technology

1. **Introduction.** Bounds in the mean square sense for elastostatic boundary value problems have been found by Prager and Synge [1]† by a method in which they split the original boundary value problem into two relaxed problems such that the common solution of the two relaxed problems is the solution of the original problem. They associate a set of six stress components with a point or vector in function space, and define the metric of the function space by a suitable choice of the scalar product. They show that the solution of the original boundary value problem, interpreted as a vector in function space, must lie on a hypercircle. Since the metric of the function space is positive definite, the hypercircle is bounded, so that the location of the solution of the original boundary value problem on the hypercircle bounds it in the mean square sense. They further show that these bounds may be improved by introducing additional solutions of the relaxed problems.

In a later paper by Synge [2], the method of the hypercircle is applied to boundary value problems other than the elastostatic problems; these include the Dirichlet and Neumann problems. In that paper bounds are derived for the mean value of the solution in the neighborhood of a point. The method used differs from that of the present paper, but there are certain underlying ideas in common.

In the present paper, we consider only the Dirichlet problem; we assume that the solution has already been located on a hypercircle in function space, using the general approach of the two preceding papers; we then proceed to determine bounds for the solution at a point.

We consider the problem in Euclidean N-space E_N . Apart from any further significance of this generalization, it enables us to discuss simultaneously the two most interesting cases, $N = 2$, $N = 3$.

In addition to determining bounds for the solution at a point, the same general method makes it possible to obtain bounds for the various derivatives of the solution. Bounds for the derivatives are considered separately in Sec. 4.

The work up to this point is concerned specifically with the problem of determining bounds for the solution and its derivatives at a point interior to the domain of definition

*Received Dec. 10, 1948. The material contained in this paper comes from a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Science at Carnegie Institute of Technology.

**Now at Iowa State College.

†Numbers in brackets refer to the bibliography at the end of the paper.

of our problem. However, for a certain type of problem (e.g. the torsion problem), it is useful to have bounds for the derivatives at points on the boundary. In Sec. 5, bounds are obtained for the normal derivative at a point on a plane face.

In Sec. 6, a numerical example is given to illustrate the practical application of the method. At a specified point, two bounds are obtained for the numerical value of the solution of the Dirichlet problem in a plane region.

An alternative method, suggested by the work of Greenberg, is discussed in Sec. 7. This method is discussed for the two dimensional problem, but the reader may extend it to N-dimensions without difficulty.

2. Notation and auxiliary theorem. Latin suffixes take the range 1, 2, $\cdots$, N, and the summation convention holds for repeated suffixes. The summation convention does not hold for Greek suffixes. The coordinates are rectangular cartesian, and differentiation with respect to a coordinate is indicated by a subscript following a comma ($u_{,i} = \partial u/\partial x_i$).

We shall denote by V an open domain in Euclidean N-space E_N, bounded by a closed surface B of $(N - 1)$ dimensions: V may be simply or multiply connected. We shall have occasion to make use of an N-dimensional sphere with center at a general point P (contained in V) and with radius a. The N-dimensional interior of any such sphere will be denoted by v and its bounding $(N - 1)$-surface by b (Fig. 1). The unit normal

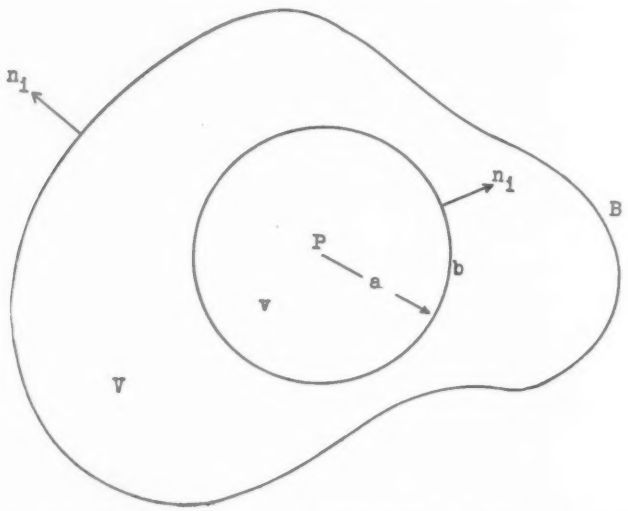

FIG. 1. In Euclidean N-space, V is an arbitrary open domain bounded by an $(N - 1)$-dimensional surface B. The point P is surrounded by an N-dimensional sphere b of radius a and center at P: v is the volume enclosed by b. The unit normal n_i is directed outward from V on B and outward from v on b.

$n_i (i = 1, 2, \cdots, N)$ will always be directed outward from V on B and outward from v on b.

In all cases, integration will be denoted by a single sign and the range of integration indicated by the element of integration unless otherwise stated.

We introduce the idea of function space, which we shall denote by F, taking care to distinguish between a vector in F and a vector in the Euclidean N-space, E_N. The former will be denoted by heavy type, $\mathbf{S}$; the latter will carry a subscript, u_i.

We define a vector $\mathbf{S}$ in F to be any vector field p_i in $V + B$. We require p_i to satisfy certain differentiability conditions, namely, that V can be divided into a finite number of subregions such that p_i is continuous and has continuous first derivatives in each such subregion. We further require that the normal component $(p_i n_i)$ of p_i be continuous across the $(N - 1)$-dimensional surface separating two such subregions.

The scalar product of two vectors $\mathbf{S}$ and $\mathbf{S}'$ will be denoted by $\mathbf{S}\cdot\mathbf{S}'$ and defined by

$$\mathbf{S}\cdot\mathbf{S}' = \int p_i p_i' \, dV, \tag{2.1}$$

where the volume element dV indicates that the integration extends throughout V. On putting $\mathbf{S}' = \mathbf{S}$, we get

$$|\,\mathbf{S}\,|^2 = \mathbf{S}^2 = \mathbf{S}\cdot\mathbf{S} = \int p_i p_i \, dV. \tag{2.2}$$

This gives the metric in F and $\mathbf{S} = 0$ implies $p_i = 0$ throughout V.

Synge [2] has established the fact that the extremity of the solution vector $\mathbf{S}$ lies on a hypercircle Γ in F and has expressed this fact by writing

$$\mathbf{S} = \mathbf{C} + R\mathbf{J}, \tag{2.3}$$

where the extremity of $\mathbf{C}$ is the center of Γ and R its radius: $\mathbf{C}$ and R are considered to be known quantities for present purposes and are given explicitly by the following formulas:

$$\begin{aligned} R^2 &= \frac{1}{4}\left[\mathbf{S}^{*2} - \sum_{\mu=1}^{m_1} (\mathbf{S}^*\cdot\mathbf{I}_\mu')^2 - \sum_{\nu=1}^{m_2} (\mathbf{S}^*\cdot\mathbf{I}_\nu'')^2\right], \\ \mathbf{C} &= \frac{1}{2}\left[\mathbf{S}^* - \sum_{\mu=1}^{m_1} \mathbf{I}_\mu'(\mathbf{S}^*\cdot\mathbf{I}_\mu') + \sum_{\nu=1}^{m_2} \mathbf{I}_\nu''(\mathbf{S}^*\cdot\mathbf{I}_\nu'')\right]. \end{aligned} \tag{2.4}$$

The vector $\mathbf{S}^*$ is a known vector, and the vectors $\mathbf{I}_\mu'$ and $\mathbf{I}_\nu''$ are two orthonormal sets of unit vectors which together form an orthonormal set of $(m_1 + m_2)$ unit vectors, $\mathbf{I}_\rho$:

$$\mathbf{I}_\rho\cdot\mathbf{I}_\sigma = \delta_{\rho\sigma}, \qquad (\rho, \sigma = 1, 2, \cdots, m = m_1 + m_2), \tag{2.5}$$

where $\delta_{\rho\sigma}$ is the Kronecker delta. The vector $\mathbf{J}$ is arbitrary except for the conditions expressing that $\mathbf{J}$ is a unit vector, and that it is orthogonal to the known orthonormal set $\mathbf{I}_\rho$:

$$\mathbf{J}\cdot\mathbf{J} = 1, \qquad \mathbf{J}\cdot\mathbf{I}_\rho = 0, \qquad (\rho = 1, 2, \cdots, m). \tag{2.6}$$

Let $\mathbf{G}$ be any vector in F and consider the scalar product $\mathbf{S}\cdot\mathbf{G}$ of $\mathbf{G}$ and the solution vector $\mathbf{S}$. We determine maximum and minimum values for $\mathbf{S}\cdot\mathbf{G}$ as $\mathbf{S}$ ranges over the hypercircle Γ.

We may represent the vector $\mathbf{G}$ as the sum of its projections on each of the unit vectors $\mathbf{I}_\rho$ $(\rho = 1, 2, \cdots, m)$ and on the subspace of the hypercircle: that is

$$\mathbf{G} = M\mathbf{J}_0 + \sum_{\rho=1}^{m} N_\rho \mathbf{I}_\rho, \tag{2.7}$$

where $M\mathbf{J}_0$ ($\mathbf{J}_0$ is a unit vector) is the vector projection of $\mathbf{G}$ on the plane of the hypercircle and

$$N_\rho = \mathbf{G}\cdot\mathbf{I}_\rho. \tag{2.8}$$

To obtain M we square (2.7) and again use the orthogonality conditions; we obtain

$$M^2 = \mathbf{G}^2 - \sum_{\rho=1}^{m} (\mathbf{G}\cdot\mathbf{I}_\rho)^2. \tag{2.9}$$

This value is always positive, so that M is real, and we shall understand by M the positive square root of (2.9). Making use of (2.3), we may write

$$\mathbf{S}\cdot\mathbf{G} - \mathbf{C}\cdot\mathbf{G} = R\ \mathbf{J}\cdot\mathbf{G} = R\ \mathbf{J}\cdot\left[M\,\mathbf{J}_0 + \sum_{\rho=1}^{m} N_\rho \mathbf{I}_\rho\right] = RM\ \mathbf{J}\cdot\mathbf{J}_0\ . \tag{2.10}$$

Hence

$$|\ \mathbf{S}\cdot\mathbf{G} - \mathbf{C}\cdot\mathbf{G}\ | = RM\ |\ \mathbf{J}\cdot\mathbf{J}_0\ | \leq RM. \tag{2.11}$$

This may be summed up in the following theorem.

Theorem I. Let $\mathbf{G}$ *be an arbitrary vector in* F *and* $\mathbf{S}$ *the solution vector, the extremity of which is confined to lie on a hypercircle in* F. *Then bounds for the scalar product* $\mathbf{S}\cdot\mathbf{G}$ *are given by*

$$|\ \mathbf{S}\cdot\mathbf{G} - \mathbf{C}\cdot\mathbf{G}\ | \leq RM, \tag{2.12}$$

Where R *is the radius of the hypercircle and* M *is the positive square root of* (2.9).

3. Bounds for the solution at an interior point. The Dirichlet problem requires the determination of a function ϕ which is harmonic in a volume V and assumes assigned values on its surface B. These conditions may be written

$$\Delta\phi = 0, \qquad (\phi)_B = f, \tag{3.1}$$

where Δ is the Laplacian operator in E_N and f is a function assigned on the $(N-1)$-dimensional bounding surface B; it will be assumed that f is piecewise continuous. The normal n_i $(i = 1, 2, \cdots, N)$ to B will be assumed to be piecewise continuous and such that a solution ϕ exists which makes the various integrals occurring converge.

It is well known that the fundamental solution of the Laplace equation is

$$\begin{aligned} G &= \ln r\ , && \text{for } N = 2, \\ &= \frac{r^{2-N}}{(2-N)}, && \text{for } N \geq 3. \end{aligned} \tag{3.2}$$

As a first step toward obtaining bounds for the solution ϕ at a point P with coordinates $y_i = c_i$, we define in V a free Green's vector, obtained by differentiation of the fundamental solution:

$$\begin{aligned} G_i &= x_i r^{-N} && \text{for } r \geq a, \qquad (i = 1, 2, \cdots, N) \\ &= 0 && \text{for } r < a. \end{aligned} \tag{3.3}$$

This formula applies for $N = 2, 3, \cdots$. The coordinates x_i are relative to $P(x_i = y_i - c_i)$, r is the distance measured from $P(r^2 = x_i x_i)$, and a does not exceed the distance from P to the nearest point of B, so that the sphere $r = a$ with center at P does not cut B. It should be noted that G_i defined by (3.3) is divergence-free, that is,

$$G_{i,i} = Nr^{-N} - Nx_i x_i r^{-N-2} = 0, \tag{3.4}$$

since $x_i x_i = r^2$. This property is also evident from the fact that G_i is the gradient of the fundamental harmonic function.

We denote by $\mathbf{G}$ the vector in F corresponding to G_i in V and associate the vector $\mathbf{S}$ in F with the vector field $\phi_{,i}$ in V, where ϕ is the solution of (3.1). We seek to bound ϕ at the point P. Noting that G_i has no singularity in V, the scalar product $\mathbf{S}\cdot\mathbf{G}$ is evaluated by use of Green's theorem:

$$\begin{aligned}\mathbf{S}\cdot\mathbf{G} &= \int \phi_{,i}G_i\,dV = \int_{V-v} \phi_{,i}G_i\,dV \\ &= \int \phi G_i n_i\,dB - \int \phi G_i n_i\,db - \int_{V-v} \phi G_{i,i}\,dV,\end{aligned} \tag{3.5}$$

where db is an element of the sphere ($r = a$). The first surface integral is calculable since ϕ is known on B. The normal n_i points outward from v on b and is equal to x_i/r. Then by use of a mean value theorem for harmonic functions[1], the integral over b is expressible in terms of the value of the function ϕ at the point P:

$$\begin{aligned}\int \phi G_i n_i\,db &= \int \phi n_i x_i r^{-N}\,db = \int \phi r^{1-N}\,db \\ &= a^{1-N}\int \phi\,db = K_N\phi(P),\end{aligned} \tag{3.6}$$

where K_N is a constant depending only on the dimensionality of our space and is given explicitly by (3.9).

We may now write (3.5) in the form:

$$K_N\phi(P) = \int \phi G_i n_i\,dB - \mathbf{S}\cdot\mathbf{G}. \tag{3.7}$$

This relation expresses the value of the solution at the point P in terms of a calculable quantity and the scalar product of the solution vector and the Green's vector in the function space F.

Theorem I of Section 2 provides bounds for the scalar product; so combining these results, we have bounds for the solution at the point P which are given by the following theorem.

Theorem II. Let ϕ be a function which is harmonic in V and assumes assigned values on the boundary B of V. Let P be any interior point in V. Then bounds for $\phi(P)$ are given by the inequality

$$\left| \int f G_i n_i\,dB - \mathbf{C}\cdot\mathbf{G} - K_N\phi(P) \right| \leq MR, \tag{3.8}$$

where n_i is the unit normal outward from V on B.

In (3.8), $\mathbf{C}$, R are given by (2.4), M is the positive square root of (2.9), G_i is given by (3.3), and

$$\begin{aligned} K_N &= \frac{2\pi^{N/2}}{\Gamma(N/2)}, \qquad (K_2 = 2\pi,\ K_3 = 4\pi), \\ f &= (\phi)_B\,, \\ \mathbf{C}\cdot\mathbf{G} &= \int_{V-v} C_i x_i r^{-N}\,dV, \end{aligned} \tag{3.9}$$

[1] O. D. Kellogg, *Foundations of potential theory*, Berlin, 1929, pp. 223-224.

where C_i denotes the vector field in V corresponding to the vector $\mathbf{C}$ in F whose extremity lies at the center of the hypercircle.

4. Bounds for the derivatives of the solution at an interior point. To obtain bounds for the first derivatives, we define a new Green's vector $G_i^{(p)}$ which is derived from the Green's vector of Section 3 by differentiation of equation (3.3) with respect to x_p :

$$\begin{aligned} G_i^{(p)} &= \delta_{ip} r^{-N} - N x_i x_p r^{-N-2}, \qquad \text{for } r \geq a, \\ &= 0, \qquad\qquad\qquad\qquad \text{for } r < a, \qquad (i, p = 1, 2, \cdots, N). \end{aligned} \tag{4.1}$$

It is of course obvious that this vector is divergence-free since it is the second derivative of the fundamental harmonic function.

We denote by $\mathbf{G}^{(p)}$ the vector in the function space F corresponding to $G_i^{(p)}$ in V and evaluate the scalar product $\mathbf{S}\cdot\mathbf{G}^{(p)}$, making use of Green's theorem:

$$\begin{aligned} \mathbf{S}\cdot\mathbf{G}^{(p)} &= \int \phi_{,i} G_i^{(p)}\, dV = \int_{V-v} \phi_{,i} G_i^{(p)}\, dV \\ &= \int \phi G_i^{(p)} n_i\, dB - \int \phi G_i^{(p)} n_i\, db - \int_{V-v} \phi G_{i,i}^{(p)}\, dV. \end{aligned} \tag{4.2}$$

The integral over B is calculable since ϕ is known on B, and the last integral vanishes since $G_i^{(p)}$ is divergence-free. We evaluate the remaining integral, making use of Green's theorem and another mean value theorem for harmonic functions[2]. We observe that $n_i = x_i/r$ and $n_i x_p = n_p x_i$ on b. Then

$$\begin{aligned} \int \phi G_i^{(p)} n_i\, db &= a^{-N} \int \phi n_p\, db - N a^{-N-2} \int \phi x_i x_p n_i\, db \\ &= (1 - N) a^{-N} \int \phi_{,p}\, dv = (1 - N) L_N\, \phi_{,p}(P), \end{aligned} \tag{4.3}$$

where L_N depends only on the dimensionality of the space and is given explicitly by

$$L_N = K_N/N = \frac{\pi^{N/2}}{\Gamma(N/2 + 1)}. \tag{4.4}$$

Equation (4.2) is now in the form:

$$(1 - N) L_N \phi_{,p}(P) = \int \phi G_i^{(p)} n_i\, dB - \mathbf{S}\cdot\mathbf{G}^{(p)}. \tag{4.5}$$

This relation gives an expression for the value of the first derivative of the solution at the point P in terms of a calculable integral and the scalar product of the solution vector and a Green's vector in the function space F.

The result (4.5) when combined with Theorem I gives bounds for the first derivatives of the solution ϕ at the point P. Thus, we have established the following theorem.

Theorem III. Let ϕ be a function which is harmonic in V and assumes assigned values on the boundary B of V. Then bounds for the first derivatives of the solution at a point P contained in V are given by the inequality

[2] *Loc. cit.* pp. 223-224.

$$\left| \int f G_i^{(p)} n_i \, dB - \mathbf{C} \cdot \mathbf{G}^{(p)} + (N-1) L_N \phi_{,p}(P) \right| \leq M^{(p)} R, \tag{4.6}$$

where n_i represents the unit normal outward from V on B.

The value of L_N is given by (4.4) and $M^{(p)}$ is obtained from (2.9) by replacing $\mathbf{G}$ by $\mathbf{G}^{(p)}$.

A process similar to the preceding gives bounds for the higher order derivatives of ϕ at a point P interior to V. The appropriate vector field in V is obtained in each case by differentiation of (3.3). The variables with respect to which (3.3) is to be differentiated are determined by the particular derivative of ϕ for which bounds are being sought. For example, if bounds are required for $\phi_{,pq}$, then the vector field is obtained for this case by differentiation of (3.3) with respect to x_p and x_q. The resulting vector field is then made to correspond to a Green's vector in the function space F and the scalar product of this Green's vector and the solution vector is computed. Combining the results of this computation with Theorem I gives bounds for the derivative under consideration. Such results may be stated in a manner analogous to Theorem III with appropriate changes. Thus, if bounds were required for $\phi_{,pq}$ the inequality (4.6) must be replaced by

$$\left| \int f G_i^{(pq)} n_i \, dB - \mathbf{C} \cdot \mathbf{G}^{(pq)} - \frac{N^2}{N+2} L_N \phi_{,pq}(P) \right| \leq M^{(pq)} R, \tag{4.7}$$

with the obvious meaning given to the changed notation.

5. Bounds for the normal derivative at a point on a plane boundary. The preceding work is concerned with the problem of determining bounds for the solution and its

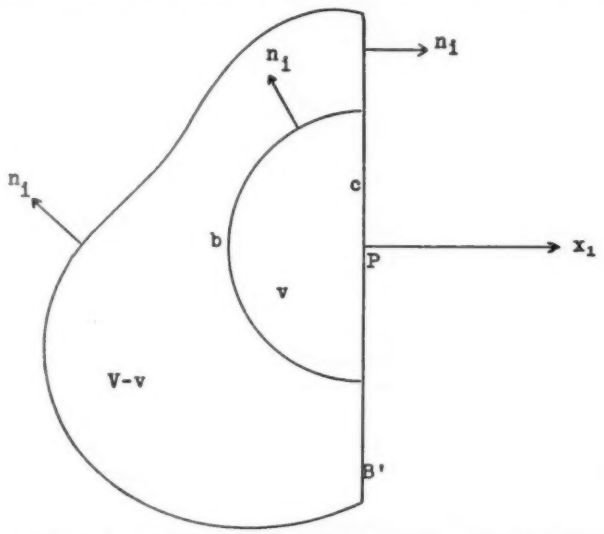

FIG. 2. $\Delta\phi = 0$ in V, $(\phi)_B = f$, a given function. The origin of coordinates is at P and the x_1-axis is perpendicular to the plane portion of the bounding surface B. The hemispherical domain surrounding P is bounded by $b + c$, where $c = B - B'$.

derivatives at a point interior to the domain V. However in certain types of problems, it is useful to have bounds for the normal derivative at a point on the boundary. In this

section, we confine our attention to a domain V bounded by a closed surface B such that at least one portion of B is a plane (Fig. 2). We shall establish bounds for the normal derivative of the solution at a point P on this plane portion of the bounding surface.

A system of rectangular cartesian coordinates is introduced such that the origin is at P and the positive x_1-axis is perpendicular to B and is directed outward from V. Thus the derivative $\phi_{,1}$ is the normal derivative of ϕ at P. The point P is then surrounded by a hemispherical domain v of radius a lying entirely in V. The surface of the hemisphere is denoted by b and that portion of B which caps b is called c. We use B' to denote $B - c$.

We define a Green's vector $\mathbf{G}^{(1)}$ in the function space F corresponding to the vector field in V obtained from (4.1) by setting $p = 1$, and compute the scalar product of $\mathbf{G}^{(1)}$ and the solution vector $\mathbf{S}$. However after applying Green's theorem, the surface integrals over B and the boundary of v have the range of integration c in common so that the scalar product takes the form

$$\mathbf{S}\cdot\mathbf{G}^{(1)} = \int_{B'} \phi G_i^{(1)} n_i \, dB' - \int_b \phi G_i^{(1)} n_i \, db. \tag{5.1}$$

As in the previous instances, the first integral is calculable since $\phi = f$ on B', but we no longer have the mean value property of harmonic functions to aid in the evaluation of the last integral since it is an integral over the hemispherical surface b. However, this last integral can be evaluated by other means to obtain the result

$$\int_b \phi G_i^{(1)} n_i \, db = a^{-N} \int_0^a z^{N-1} \, dz \int_0^z F(x) \, dx + (N-1)a^{-N} \int_c f \, dc + \frac{1}{2}(1-N)L_N \phi_{,1}(P), \tag{5.2}$$

where

$$F(x) = (N-1)x^{1-N} \int_c (\phi_{,22} + \phi_{,33} + \cdots + \phi_{,NN}) \, dc. \tag{5.3}$$

Since $\phi = f$ on c and the x_1-axis is perpendicular to c, the integrand of (5.3) is a known function, hence F is a known function. We note that this evaluation of the last integral of (5.1) requires that the boundary function f and its first and second derivatives be piecewise continuous and such that their integrals exist.

Upon substitution from (5.2) into (5.1), an explicit expression for $\mathbf{S}\cdot\mathbf{G}^{(1)}$ is obtained in terms of calculable integrals and the normal derivative of ϕ at the point P. Then combining this result with Theorem I, we obtain the following theorem.

Theorem IV. Let ϕ be a function which is harmonic in V and assumes assigned values on the boundary B of V. It is further required that B contains a plane portion. Then bounds for the normal derivative of ϕ at a point P on this plane portion are given by the inequality

$$\left| \int_{B'} f G_i^{(1)} n_i \, dB' - \mathbf{C}\cdot\mathbf{G}^{(1)} - a^{-N} \int_0^a z^{N-1} \, dz \int_0^z F(x) \, dx - (N-1)a^{-N} \int_c f \, dc - \frac{1}{2}(1-N)L_N \frac{\partial \phi(P)}{\partial n} \right| \leq M^{(1)} R, \tag{5.4}$$

where n_i is the unit outward normal.

It is interesting to note that Theorem IV gives an exact value in the following case. Let ϕ be harmonic in the domain V, where V is a volume in Euclidean N-space bounded by a hemispherical surface B' of radius a and a diametral plane c. The point P is taken at the center of curvature of B' (Fig. 3). We assign ϕ the value zero on c and unity on B'

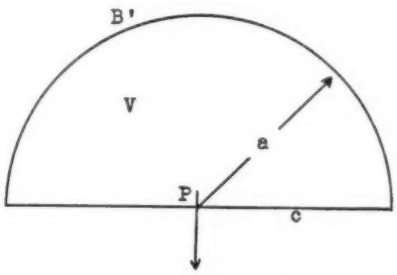

FIG. 3. $\Delta\phi = 0$ in V, $(\phi)_{B'} = 1$, $(\phi)_c = 0$. The normal derivative of ϕ at the point P is computed for the special case in which V consists of a domain bounded by a hemispherical surface and a diametral plane in Euclidean N-space.

and seek the value of the normal derivative of ϕ at the point P. We may take the volume v to coincide with V, so that obviously $M^{(1)}$ and $\mathbf{C}\cdot\mathbf{G}^{(1)}$ of (5.4) are zero. Further, since $\phi = 0$ on c, it follows that $F(a) = 0$ and that the integral of f over c is also zero. The remaining integral over B', when evaluated, gives

$$\int_{B'} fG_i^{(1)} n_i \, dB' = \frac{(1 - N)\pi^{(N-1)/2}}{a\Gamma[(N - 1)/2]} . \tag{5.5}$$

Substitution of (5.5) into (5.4) gives

$$\phi_{,1}(P) = \frac{(N - 1)L_{N-1}}{a\, L_N} . \tag{5.6}$$

For $N = 2$, $N = 3$, this gives respectively

$$\begin{aligned} \phi_{,1}(P) &= 2/(a\pi), \qquad (N = 2), \\ \phi_{,1}(P) &= 3/(2a), \qquad (N = 3). \end{aligned} \tag{5.7}$$

This quantity represents the intensity at the center of a circular ($N = 2$) or spherical ($N = 3$) condenser which is split symmetrically and the two halves raised to potentials $+1$ and -1.

6. **A numerical example.** As an illustration of the use of the method contained in the preceding sections, we consider the following problem: let the open domain V be the area in the Euclidean plane ($N = 2$) bounded by B, $x = \pm 1$, $x^2 + y^2 = 2$ (Fig. 4). We seek bounds for the function ψ which satisfies the conditions

$$\Delta\psi = 0, \qquad (\psi)_B = f = \tfrac{1}{2}(x^2 + y^2). \tag{6.1}$$

In particular, we establish numerical bounds for the function ψ at the origin (the point P in the notation of the preceding work).

We use the explicit (x, y) notation and the implicit notation interchangeably, depending on which is the more convenient.

We first find a vector field p_i^* and a scalar field ψ^* in V satisfying the conditions

$$p_i^* = \psi^*_{.i} , \qquad (\psi^*)_B = f. \tag{6.2}$$

Secondly, we find a vector field p_i'' in V satisfying

$$p''_{i,i} = 0 \tag{6.3}$$

and no boundary conditions. The common solution of (6.2) and (6.3) is the solution of (6.1). We can obtain solutions of (6.3) by means of skew differentiation: $p_i'' = \epsilon_{ij}\chi''_{.j}$.

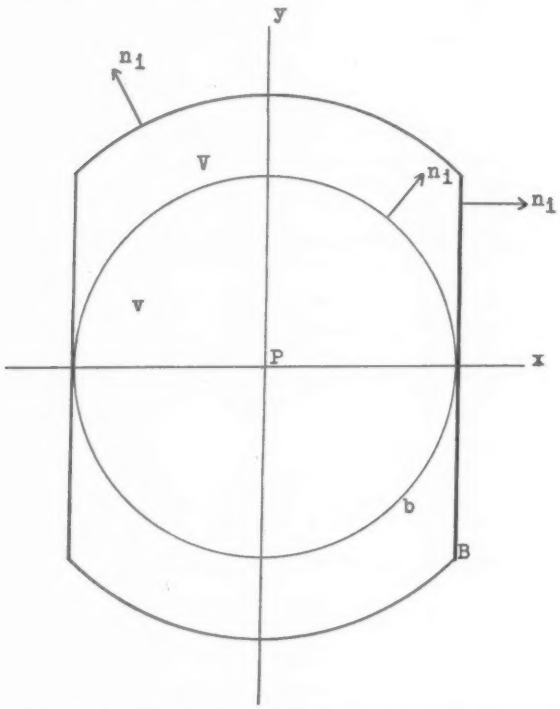

FIG. 4. $\Delta\psi = 0$ in V, $(\psi)_B = \frac{1}{2}(x^2 + y^2)$. The domain V is located in the Euclidean plane and is bounded by B: $(x^2 + y^2 - 2)(x^2 - 1) = 0$. The point P is taken at the origin and v consists of the interior of the unit circle.

For the purpose of improving the bounds, we use a set of vectors of the class (6.3) and introduce another class of vector fields p_i' in V satisfying the conditions

$$p_i' = \psi'_{.i} , \qquad (\psi')_B = 0. \tag{6.4}$$

Thus our problem is that of selection of the three types of functions from which we can derive the vector fields p_i^* , p_i'', p_i' satisfying the appropriate conditions. The functions selected for use in our particular example are given in Table I. This table shows the

vector in F, the vector field in V to which it corresponds, and the particular function from which the vector field is derived. The functions given in Table I are used to obtain only two approximations for the solution function ψ. It is, of course, possible to improve the approximation last obtained by the selection of additional functions.

TABLE I.

Vector in F	Vector field in V	Function from which vector field is derived
$\mathbf{S}^*$	$p_1^* = x$ $p_2^* = y$	$\psi^* = (x^2 + y^2)/2$
$\mathbf{S}_1'$	$p'_{(1)1} = 2x(2x^2 + y^2 - 3)$ $p'_{(1)2} = 2y(x^2 - 1)$	$\psi'_{(1)} = (x^2 - 1)(x^2 + y^2 - 2)$
$\mathbf{S}_2'$	$p'_{(2)1} = 2x(3x^4 + 2x^2y^2 - 6x^2 - y^2 + 2)$ $p'_{(2)2} = 2x^2y(x^2 - 1)$	$\psi'_{(2)} = x^2(x^2 - 1)(x^2 + y^2 - 2)$
$\mathbf{S}_3'$	$p'_{(3)1} = 2xy^2(2x^2 + y^2 - 3)$ $p'_{(3)2} = 2y(x^2 - 1)(x^2 + 2y^2 - 2)$	$\psi'_{(3)} = y^2(x^2 - 1)(x^2 + y^2 - 2)$
$\mathbf{S}_1''$	$p''_{(1)1} = x$ $p''_{(1)2} = -y$	$\chi''_{(1)} = xy$
$\mathbf{S}_2''$	$p''_{(2)1} = x^3 - 3xy^2$ $p''_{(2)2} = y^3 - 3x^2y$	$\chi''_{(2)} = x^3y - xy^3$

To determine what computations are required, we examine the quantities contained in the inequality (3.8) of Theorem II, Section 3; namely, for $N = 2$,

$$\left| \int fG_i n_i \, dB - \mathbf{C}\cdot\mathbf{G} - 2\pi\psi(P) \right| \leq MR. \tag{6.5}$$

It is convenient to have all the scalar products available before computing the various quantities present in this inequality. These scalar products have been tabulated in a triangular array in Table II. The number in Table II which represents the scalar product of any two vectors $\mathbf{S}$, $\mathbf{S}'$ is located at the intersection of row $\mathbf{S}$ and column $\mathbf{S}'$ (or vice versa).

TABLE II. Scalar products.

$\mathbf{S}^*$	4.474926						
$\mathbf{S}_1'$	−8.633233	21.474728					
$\mathbf{S}_2'$	−1.412487	1.458289	3.253382				
$\mathbf{S}_3'$	−3.170796	6.953487	.395429	11.049406			
$\mathbf{S}_1''$	−1.333333	0	0	0	4.474926		
$\mathbf{S}_2''$	− .533333	0	0	0	−2.133333	7.654613	
$\mathbf{G}$	2.000000	−3.141593	−.785398	−2.356194	− .858407	−.050148	1.493808
	$\mathbf{S}^*$	$\mathbf{S}_1'$	$\mathbf{S}_2'$	$\mathbf{S}_3'$	$\mathbf{S}_1''$	$\mathbf{S}_2''$	$\mathbf{G}$

We next orthonormalize the vectors $\mathbf{S}_i'$, $\mathbf{S}_i''$. It is inherent in the method that $\mathbf{S}_i' \cdot \mathbf{S}_j'' = 0$ for all i, j (this may easily be checked by integration), so it is only necessary to require that

$$\begin{aligned} \mathbf{S}_i' \cdot \mathbf{S}_j' &= \delta_{ij} , \\ \mathbf{S}_i'' \cdot \mathbf{S}_j'' &= \delta_{ij} . \end{aligned} \tag{6.6}$$

A systematic method of orthonormalization has been given by Peach [7]. We apply his method to the two preceding classes of vectors $\mathbf{S}_i'$, $\mathbf{S}_i''$ to obtain the two orthonormal

TABLE III. Scalar products.

	$\mathbf{I}_1'$	$\mathbf{I}_2'$	$\mathbf{I}_3'$	$\mathbf{I}_1''$	$\mathbf{I}_2''$
$\mathbf{S}^*$	−1.863019	.465193	−.133665	−.630298	.453794
$\mathbf{G}$	− .677944	.322103	−.456269	−.405789	.178329

$$(\mathbf{S}^* \cdot \mathbf{I}_1')^2 = 3.470839 \qquad (\mathbf{G} \cdot \mathbf{I}_1')^2 = .459608$$

$$(\mathbf{S}^* \cdot \mathbf{I}_2')^2 = .216405 \qquad (\mathbf{G} \cdot \mathbf{I}_2')^2 = .103750$$

$$(\mathbf{S}^* \cdot \mathbf{I}_3')^2 = .017866 \qquad (\mathbf{G} \cdot \mathbf{I}_3')^2 = .208181$$

$$(\mathbf{S}^* \cdot \mathbf{I}_1'')^2 = .397275 \qquad (\mathbf{G} \cdot \mathbf{I}_1'')^2 = .164665$$

$$(\mathbf{S}^* \cdot \mathbf{I}_2'')^2 = .205929 \qquad (\mathbf{G} \cdot \mathbf{I}_2'')^2 = .031801$$

classes $\mathbf{I}_i'$, $\mathbf{I}_i''$ (obviously, $\mathbf{I}_i' \cdot \mathbf{I}_j'' = 0$, for all i, j). When these computations have been carried out, we find:

$$\begin{aligned}
\mathbf{I}_1' &= .215792\ \mathbf{S}_1' , \\
\mathbf{I}_2' &= -.038235\ \mathbf{S}_1' + .563047\ \mathbf{S}_2' , \\
\mathbf{I}_3' &= -.109735\ \mathbf{S}_1' + .008205\ \mathbf{S}_2' + .337178\ \mathbf{S}_3' , \\
\mathbf{I}_1'' &= .472723\ \mathbf{S}_1'' , \\
\mathbf{I}_2'' &= .185041\ \mathbf{S}_1'' + .388146\ \mathbf{S}_2'' .
\end{aligned} \tag{6.7}$$

We also need the scalar products of $\mathbf{S}^*$ with each $\mathbf{I}'$, $\mathbf{I}''$, and the square of all such products. These are easily obtained from (6.7) and Table II: the results are given in Table III.

The approximations for R^2 are given by direct substitution from Table III into the first of (2.4). The first two approximations for R^2 are as follows:

$$\begin{aligned}
R_1^2 &= .151703, \quad \text{(approximation based on } \mathbf{S}^*, \mathbf{I}_1' , \mathbf{I}_1''), \\
R_2^2 &= .041653, \quad \text{(approximation based on all vectors given above).}
\end{aligned} \tag{6.8}$$

It should be pointed out that the above quantities are independent of the position of the point at which we are seeking to bound ψ. In fact, both the radius of the hypercircle and the position of its center depend only on the domain V and on the functions from which the vector fields are obtained. Consequently, we would write down the first two approximations for $\mathbf{C}$ if we so desired. However, we are not directly interested in $\mathbf{C}$, so we proceed to the computation of $\mathbf{C} \cdot \mathbf{G}$ and M.

We use the Green's vector defined by (3.3) for $N = 2$ and $a = 1$; that is,

$$\left.\begin{aligned} G_1 &= xr^{-2} \\ G_2 &= yr^{-2} \end{aligned}\right\} \text{ for } r \geq 1, \qquad G_1 = G_2 = 0, \quad \text{for } r < 1. \tag{6.9}$$

The necessary scalar products of $\mathbf{G}$ with each of the other vectors have been computed and are included in the appropriate tables. We then form the scalar product $\mathbf{C} \cdot \mathbf{G}$ by dotting $\mathbf{G}$ into the second of (2.4). Then direct substitution from Tables II and III gives the first two numerical approximations for $\mathbf{C} \cdot \mathbf{G}$. Substitution from Table III into (2.9) gives the approximations for M^2. These results are as follows:

$$\begin{aligned}
M_1^2 &= .869535, \quad M_1 = .932488, \quad \text{(first approximation)}, \\
\mathbf{G} \cdot \mathbf{C}_1 &= .496371, \quad \text{(first approximation)}, \\
M_2^2 &= .525803, \quad M_2 = .725122, \quad \text{(second approximation)}, \\
\mathbf{G} \cdot \mathbf{C}_2 &= .431422, \quad \text{(second approximation)}.
\end{aligned} \tag{6.10}$$

The line integral over the boundary B of the domain V is found to be

$$\int f G_{,i} n_i \, dB = 2 + \pi. \tag{6.11}$$

We may now substitute from (6.11), (6.10), (6.8) into (6.5) to obtain bounds for the function ψ at the point $P(0, 0)$. This gives

$$\begin{aligned} .682 \leq \psi(P) \leq .797, \\ .726 \leq \psi(P) \leq .773, \end{aligned} \tag{6.12}$$

as the first and second approximations respectively.

7. **An alternative method.** The reading of the work of H. J. Greenberg [8] on the same subject suggested an alternative method of attack on the problem of obtaining bounds at an interior point for the solution of the Dirichlet problem. The method is outlined below for the special case $N = 2$, but may obviously be extended to the N-dimensional case.

We let ϕ represent the solution we seek to bound at the point P and $\mathbf{S}$ the solution vector in function space corresponding to ϕ. We use two functions: the first is the free Green's function G, and the second any function H which is regular (the function and its first derivatives continuous) throughout V and is equal to G on the boundary B of V. Thus we have:

$$\begin{aligned} G &= \log r, \\ (H)_B &= (G)_B \, . \end{aligned} \tag{7.1}$$

We consider the scalar product $\mathbf{S}\cdot\mathbf{G}$ of the solution vector $\mathbf{S}$ and the Green's vector $\mathbf{G}$. We cannot apply the divergence theorem to $\mathbf{S}\cdot\mathbf{G}$ directly so we surround P with a small sphere b with center at P and radius a and remove the interior v of b. Then applying the divergence theorem to $\mathbf{S}\cdot\mathbf{G}$ for the region $V - v$ in both possible ways and making use of the fact that both ϕ and G are harmonic, we have

$$\begin{aligned} \int_{V-v} \phi_{,i} G_{,i} \, dV &= \int_B \phi G_{,i} n_i \, dB - \int_b \phi G_{,i} n_i \, db \\ &= \int_B G \phi_{,i} n_i \, dB - \int_b G \phi_{,i} n_i \, db. \end{aligned} \tag{7.2}$$

Upon taking the limit as the radius a goes to zero, we obtain the following well known expression for the value of ϕ at point P:

$$2\pi\phi(P) = \int_B \phi G_{,i} n_i \, dB - \int_B G \phi_{,i} n_i \, dB. \tag{7.3}$$

The first integral is calculable since ϕ is known on B. The essential point of the method concerns the second integral. Since the regular function H is equal to G on B, we replace G by H and again apply the divergence theorem. Since ϕ is harmonic, this gives,

$$\int_B G \phi_{,i} n_i \, dB = \int_B H \phi_{,i} n_i \, dB = \int_V H_{,i} \phi_{,i} \, dV = \mathbf{H}\cdot\mathbf{S}. \tag{7.4}$$

Thus equation (7.3) may be written:

$$2\pi\phi(P) = \int_B \phi G_{,i} n_i \, dB - \mathbf{H}\cdot\mathbf{S}. \tag{7.5}$$

Now if the extremity of $\mathbf{S}$ is confined to a known hypercircle (cf. Section 2), we know that $\mathbf{H}\cdot\mathbf{S}$ is bounded above and below. Hence (7.5) gives bounds for $\phi(P)$.

It will be recalled that in the work of Section 2 the norm of the Green's vector $|\mathbf{G}|$ enters into the bounds for $\mathbf{G}\cdot\mathbf{S}$ via the number M (see equation (2.9)); and in order to insure the existence of $|\mathbf{G}|$, it was necessary to define our vector field to be zero in a region v containing P. This difficulty does not now arise, because it is $|\mathbf{H}|$, not $|\mathbf{G}|$, that enters, and $|\mathbf{H}|$ is finite since H is regular.

The question of practical importance now is whether we can find the function H which has the required properties. This question has already been answered by Greenberg since one of his functions has these precise properties. The function referred to is defined in the following way:

$$\begin{aligned} H &= \log r, && \text{for } r \geq a, \\ &= c_1 r^2 + c_2 r + c_3, && \text{for } r < a, \end{aligned} \tag{7.6}$$

where the coefficients c_1, c_2, c_3 are determined so that H has continuous second derivatives across b, that is:

$$\begin{aligned} c_1 &= -a^{-2}/2, \\ c_2 &= 2a, \\ c_3 &= \log a - 3/2. \end{aligned} \tag{7.7}$$

But this is not the only way in which H may be defined, at least for convex regions V. For any point P interior to V, the distance r from P to any point on the boundary B may be defined by a single valued function of a polar angle θ at P; explicitly,

$$r = \rho(\theta), \qquad 0 \leq \theta \leq 2\pi. \tag{7.8}$$

Then, if we define H by the relation

$$H = \frac{r^2 \log r}{(\rho(\theta))^2} + h, \tag{7.9}$$

where h is harmonic in V and vanishes on B, we have a function such that $(H)_B = (G)_B$ and has the required regularity at P.

In conclusion, the author acknowledges his indebtedness to Professor John L. Synge under whose supervision this paper was prepared.

Bibliography

1. W. Prager and J. L. Synge, *Approximations in elasticity based on the concept of function space*, Quart. App. Math. **5**, 241-269 (1947).
2. J. L. Synge, *The method of the hypercircle in function space for boundary value problems*, Proc. Roy. Soc. (London) **A 191** 447-467 (1947).
3. O. D. Kellogg, *Foundations of potential theory*, (Berlin) (1929), 223-224.

4. R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Berlin, 1937, Vol. 2, p. 251.
5. R. Courant, *Differential and integral calculus*, translated by E. J. McShane, New York, Vol. 2, pp. 302-304.
6. P. Weiss, *An extension of Cauchy's integral formula by means of Maxwell's stress tensor*, J. London Math. Soc. **21**, 210-218 (1946).
7. M. O. Peach, *Simplified technique for constructing orthonormal functions*, Bull. Amer. Math. Soc. **50**, 556-564 (1944).
8. H. J. Greenberg, *The determination of upper and lower bounds for the solution of the Dirichlet problem*, J. Math. Phys. **27**, 161-182 (1949).

THE ELECTROMAGNETIC FIELD PRODUCED BY A HELIX*

BY

R. S. PHILLIPS

University of Southern California

1. Introduction. In this paper[1] we are concerned with finding the electromagnetic field produced by an infinitely-thin perfectly-conducting helix subject to a monochromatic excitation. Because of the use of the helix in an ultra-high frequency amplifier known as the traveling-wave tube [2], there has been in recent years considerable interest in this problem. Earlier papers dealing with the field due to a helix have been based on a model consisting of a circular cylinder which conducts only in a given helical direction.[2] [1], [2], [3]. The symmetry of this idealization corresponds to that of cylindrical coordinates so that such a model yields to the usual separation of variables technique. Our approach is quite different.

We start off by assuming that the effect of a monochromatic electromagnetic field would be to induce a sinusoidal electric current on the helical conductor; and that this current travels along the helix with a real propagation constant β in the axial direction.[3] The electromagnetic field accompanying such a current may then be represented by means of the retarded vector potential Π as [5, p. 165]

$$\Pi(x, y, z, t) = A \int ([i]/r)\, ds, \tag{1.1}$$

where $[i]$ is the current vector at a point (ξ, η, ζ) of the helix at time $(t - r/c)$, c being the free space wave velocity, $r = [(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2]^{1/2}$, s the arc length along the helix, and A a scalar factor.

It follows that

$$E = \nabla(\nabla \cdot \Pi) + k^2\Pi, \qquad H = -i\omega\epsilon\, \nabla \times \Pi, \tag{1.2}$$

where $k^2 = \epsilon\mu\omega^2$, ω being the radial frequency of the field. We then attempt to determine β so as to make the electric field along the prefectly conducting helix "essentially" zero, and thus validate our initial assumption. It will turn out that the β so determined corresponds to a phase velocity along the helix itself equal to c.

This then is our program. The principal difficulty lies in the singular nature of the potential function Π. In Sec. 2 we show that the integral defining Π converges; and in addition we obtain an evaluation for Π in terms of Bessel functions. In fact we establish

$$\int_{-\infty}^{\infty} e^{i\beta\zeta}(e^{ikR}/R)\, d\zeta = \sum_{n=-\infty}^{\infty} 4\pi\, e^{in\varphi} I_n(\Gamma_n a) K_n(\Gamma_n \rho), \qquad (\rho > a), \tag{1.3}$$

---

*Received June 20, 1949.

[1]The work on this paper was done under funds granted to New York University by the Air Force Cambridge Research Laboratories, under contract W28-097-ac-170.

[2]By helical direction we mean the direction of a vector field which in cylindrical coordinates has the components $v_\rho = 0$, $v_\varphi = a$, $v_z = \alpha$.

[3]That is, the amplitude varies as $\cos(\beta z - \omega t)$ where z is measured in the axial direction and ω is the radial frequency.

where

$$R = [\zeta^2 + a^2 + \rho^2 - 2a\rho \cos(\varphi - \zeta/\alpha)]^{1/2}$$

and

$$\Gamma_n = [(\beta - n/\alpha)^2 - k^2]^{1/2}.$$

Here, I_n and K_n are modified Bessel and Hankel functions of nth order. If for a particular n, Γ_n is imaginary, then the corresponding term in (1.3) is to be replaced by $2\pi^2 i e^{in\varphi} J_n(\gamma_n a) H_n^{(1)}(\gamma_n \rho)$, where $\gamma_n = [k^2 - (\beta - n/\alpha)^2]^{1/2}$.

In Sec. 3 we prove that Π may be differentiated either by differentiating under the integral sign or by differentiating its series expansion termwise. We then show that Π satisfies the homogeneous wave equation at all points not on the helix. Also in Sec. 3 explicit expressions for the electromagnetic field are obtained. Finally in Sec. 4 the propagation constant β is determined so that the boundary conditions for the helix are satisfied. Since the electric field is singular on the wire, β is determined so that the ratio of the electric field in the helical direction for the proper β to that of the electric field in the helical direction for any other β approaches zero as the helix is approached radially.

In cylindrical coordinates, the equation for the helix is

$$\rho = a, \qquad z = \alpha\varphi. \tag{1.4}$$

The assumed current along the helix is the real part of

$$\begin{aligned} i_\rho &= 0, \\ i_\varphi &= Ia[a^2 + \alpha^2]^{-1/2} \exp[i(\beta z - \omega t)], \\ i_z &= I\alpha[a^2 + \alpha^2]^{-1/2} \exp[i(\beta z - \omega t)]. \end{aligned} \tag{1.5}$$

The retarded potential at the point (ρ, φ, z) and the time t is the real part of

$$\begin{aligned} \Pi_\rho &= aF \int_{-\infty}^{\infty} r^{-1} \exp(i\beta\zeta + ikr) \sin(\varphi - \zeta/\alpha)\, d\zeta, \\ \Pi_\varphi &= aF \int_{-\infty}^{\infty} r^{-1} \exp(i\beta\zeta + ikr) \cos(\varphi - \zeta/\alpha)\, d\zeta, \\ \Pi_z &= \alpha F \int_{-\infty}^{\infty} r^{-1} \exp(i\beta\zeta + ikr)\, d\zeta, \end{aligned} \tag{1.6}$$

where

$$r = [(z - \zeta)^2 + a^2 + \rho^2 - 2a\rho \cos(\varphi - \zeta/\alpha)]^{1/2}$$

and

$$F = -Ie^{-i\omega t}/4\pi i\omega\epsilon\alpha.$$

If we replace sines and cosines by their exponential representations, we see that all of the integrals in (1.6) are essential of the type found in Π_z. The notation of the next section will be somewhat simplified if we make a change of variables $\zeta' = \zeta - z$ in these integrals. Then

$$\int_{-\infty}^{\infty} e^{i\beta\zeta}(e^{ikr}/r)\,d\zeta = e^{i\beta s}\int_{-\infty}^{\infty} e^{i\beta\zeta'}(e^{ikr'}/r')\,d\zeta', \tag{1.7}$$

where

$$r' = [\zeta'^2 + a^2 + \rho^2 - 2a\rho\cos[(\varphi - z/\alpha) - \zeta'/\alpha]]^{1/2}.$$

2. Evaluation of the integral. In this section we shall evaluate the integral

$$I = \int_{-\infty}^{\infty} e^{i\beta\zeta}(e^{ikR}/R)\,d\zeta, \tag{2.1}$$

where

$$R = [\zeta^2 + a^2 + \rho^2 - 2a\rho\cos(\varphi - \zeta/\alpha)]^{1/2}.$$

We restrict our consideration to real β since the integral will clearly not converge for complex β.

The usual method of handling improper integrals, namely to transform the path of integration, is not suitable here because R has an infinite set of branch points. These are the solutions of the transcendental equation $R^2 = 0$. The device which we shall use in evaluating (2.1) consists of replacing the integral along the helix by a double integral over the enveloping cylinder of the helix; this leads to a Bessel function solution.

We define

$$I(L, r) = \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re)P(r, \sigma - \zeta/\alpha), \tag{2.3}$$

where $P(r, \sigma)$ is the Poisson kernel [7, p. 51]

$$P(r, \sigma) = \frac{1}{2\pi}\frac{1 - r^2}{1 + r^2 - 2r\cos\sigma} = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} r^{|n|}e^{in\sigma}$$

and

$$\Re = [\zeta^2 + a^2 + \rho^2 - 2a\rho\cos(\varphi - \sigma)]^{1/2}$$

For $0 \leq r < 1$, $L < \infty$, $\rho > a$, the double integral is absolutely integrable and hence the order of integration may be interchanged.

By well known properties of the Poisson kernel

$$\lim_{r\to 1-}\int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re)P(r, \sigma - \zeta/\alpha) = e^{i\beta\zeta}(e^{ikR}/R)$$

uniformly for $-L \leq \zeta \leq L$. Hence

$$I = \lim_{L\to\infty}\lim_{r\to 1-} I(L, r) \tag{2.4}$$

if this limit exists.

We wish to show not only that the limit of (2.4) exists but that the order of this double limit may be interchanged. In order to accomplish this we shall prove

a) $\lim_{L\to\infty} I(L, r)$ exists for each r $(0 \leq r < 1)$

b) $\lim_{r\to 1-} I(L, r)$ exists uniformly in L for $(0 < L < \infty)$.

Since the Fourier expansion of $P(r, \sigma)$ converges uniformly in σ for each $r < 1$, it follows that we can interchange the order of integration and summation and so express $I(L, r)$ as

$$I(L, r) = \sum_{n=-\infty}^{\infty} r^{|n|} \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma \, \Re^{-1} \exp(i\beta\zeta + ik\Re) \exp[in(\sigma - \zeta/\alpha)]. \quad (2.5)$$

In order to establish (a) we first examine the individual terms of this sum. For the case $(\beta - n/\alpha) > k$, we make the substitution

$$\zeta = [a^2 + \rho^2 - 2a\rho \cos(\varphi - \sigma)]^{1/2} \sinh(\psi - \delta),$$

$$\Re = [a^2 + \rho^2 - 2a\rho \cos(\varphi - \sigma)]^{1/2} \cosh(\psi - \delta),$$

where

$$\cosh \delta = (\beta - n/\alpha)/[(\beta - n/\alpha)^2 - k^2]^{1/2},$$

$$\sinh \delta = k/[(\beta - n/\alpha)^2 - k^2]^{1/2}.$$

Then

$$(\beta - n/\alpha)\zeta + k\Re = \Phi \sinh \psi,$$

where we have set

$$\Phi = [a^2 + \rho^2 - 2a\rho \cos(\varphi - \sigma)]^{1/2}[(\beta - n/\alpha)^2 - k^2]^{1/2}. \quad (2.6)$$

Hence,

$$\int_{-L}^{L} d\zeta \, (e^{ik\Re}/\Re) \exp[i(\beta - n/\alpha)\zeta] = \int_{-l_1}^{l_2} d\psi \exp(i\Phi \sinh \psi),$$

where l_1 and l_2 are defined by

$$L = [a^2 + \rho^2 - 2a\rho \cos(\varphi - \sigma)]^{1/2} \sinh(l_1 + \delta),$$

$$L = [a^2 + \rho^2 - 2a\rho \cos(\varphi - \sigma)]^{1/2} \sinh(l_2 - \delta),$$

δ being defined as above. We now transform the path of this integral to the line $\psi = i\pi/2$. Then because of the fact that the integral along the line joining $(l, 0)$ and $(l, i\pi/2)$ goes to zero as $l \to \pm\infty$, we get

$$\int_{-\infty}^{\infty} d\zeta \, (e^{ik\Re}/\Re) \exp[i(\beta - n/\alpha)\zeta] = \int_{-\infty}^{\infty} d\psi \exp(-\Phi \cosh \psi) = 2K_0(\Phi). \quad (2.7)$$

The latter equality can be found in [6, p. 182, (7)].

It is clear from the above argument that

$$2K_0(\Phi) - \int_{-L}^{L} d\zeta \, (e^{ik\Re}/\Re) \exp[i(\beta - n/\alpha)\zeta] = \int_P d\psi \exp(i\Phi \sinh \psi),$$

where P is the sum of the two paths indicated on Fig. 1.

For $\rho > a$, the integral

$$\int_P d\psi \exp (i\Phi \sinh \psi)$$

converges to zero uniformly in σ as l_1 , $l_2 \to \infty$. This incidentally shows that the order of integration in

$$\int_0^{2\pi} d\sigma \int_{-\infty}^{\infty} d\zeta \, e^{i\beta\zeta}(e^{ik\Re}/\Re) \exp [in(\sigma - \zeta/\alpha)$$

may be interchanged. We can, however, easily obtain an upper bound in the value of this integral over P. In fact the integral from $(l, 0)$ to $(l, \pi i/2)$ is bounded by $2/(\Phi \cosh l)$;

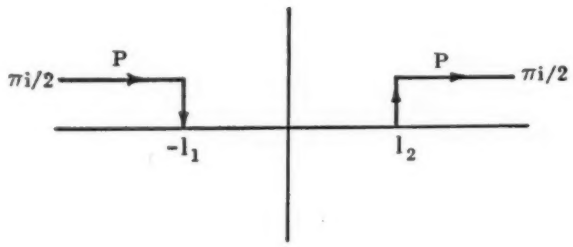

Fig. 1

whereas the integral from $(l, \pi i/2)$ to $(\infty, \pi i/2)$ is bounded by $e^{-\Phi l}/\Phi$. Now for sufficiently large $| n |$, fixed β, and fixed $\rho > a$, $\Phi \geq c | n |$. Hence in this case

$$\left| \int_0^{2\pi} d\sigma \, K_0(\Phi) \, e^{in\sigma} - \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma \, e^{i\beta\zeta}(e^{ik\Re}/\Re) \exp [in(\sigma - \zeta/\alpha)] \right| < \frac{\bar{c}}{| n |} . \qquad (2.8)$$

A similar argument shows that (2.7) and (2.8) hold equally well for $(\beta - n/2) < -k$.

For the case $(\beta - n/2)^2 < k^2$, let

$$\zeta = [a^2 + \rho^2 - 2a\rho \cos (\varphi - \sigma)]^{1/2} \sinh (\psi - \delta'),$$

$$\Re = [a^2 + \rho^2 - 2a\rho \cos (\varphi - \sigma)]^{1/2} \cosh (\psi - \delta'),$$

where now

$$\sinh \delta' = (\beta - n/\alpha)/[k^2 - (\beta - n/\alpha)^2]^{1/2},$$

$$\cosh \delta' = k/[k^2 - (\beta - n/\alpha)^2]^{1/2}.$$

For

$$\Phi' = [a^2 + \rho^2 - 2a\rho \cos (\varphi - \sigma)]^{1/2}[k^2 - (\beta - n/\alpha)^2]^{1/2}$$

we have, then,

$$(\beta - n/\alpha)\zeta + k\Re = \Phi' \cosh \psi$$

so that

$$\int_{-L}^{L} d\zeta\ (e^{ik\Re}/\Re)\ \exp\ [i(\beta - n/\alpha)\zeta] = \int_{-l'_1}^{l'_2} d\psi\ \exp\ (i\Phi' \cosh\ \psi).$$

The difference between this integral and the integral from $(-\infty - i\pi/2)$ to $(-i\pi/2)$ to $(i\pi/2)$ to $(\infty + i\pi/2)$ is the integral over the path P' shown in Fig. 2.

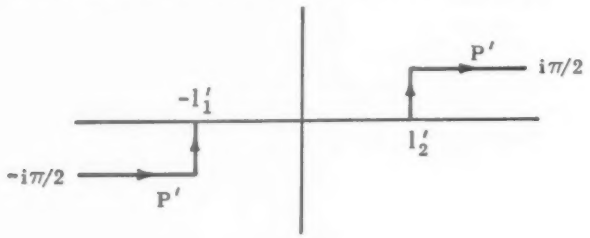

Fig. 2

For $\rho > a$, the integral $\int_{P'} d\psi\ \exp\ (i\Phi' \cosh\ \psi)$ converges to zero uniformly in σ as l_1, $l_2 \to \infty$. Hence

$$\int_{-\infty}^{\infty} d\zeta\ (e^{ik\Re}/\Re)\ \exp\ [i(\beta - n/\alpha)\zeta] = \int_{-\infty - i\pi/2}^{\infty + i\pi/2} d\psi\ \exp\ (i\Phi' \cosh\ \psi) = \pi i H_0^{(1)}(\Phi') \tag{2.9}$$

(see [6, p. 180, (8)]) and

$$\int_{-\infty}^{\infty} d\zeta \int_0^{2\pi} d\sigma\ e^{i\beta\zeta}(e^{ik\Re}/\Re)\ \exp\ [in(\sigma - \zeta/\alpha)] = \int_0^{2\pi} d\sigma\ H_0^{(1)}(\Phi')e^{in\sigma}. \tag{2.10}$$

Now by the addition theorem for Bessel functions [6, p. 361] for $\rho > a$

$$H_0^{(1)}(x[a^2 + \rho^2 - 2a\rho\ \cos\ \theta]^{1/2}) = \sum_{n=-\infty}^{\infty} J_m(xa)H_m^{(1)}(x\rho)e^{im\theta},$$

$$K_0(x[a^2 + \rho^2 - 2a\rho\ \cos\ \theta]^{1/2}) = \sum_{m=-\infty}^{\infty} I_m(xa)\ K_m(x\rho)e^{im\theta},$$

and these series converge uniformly in θ. Substituting these series in the above integrals we obtain for $\rho > a$

$$\left.\begin{aligned} &\frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta \int_0^{2\pi} d\sigma\ e^{i\beta\zeta}(e^{ik\Re}/\Re)\ \exp\ [in(\sigma - \zeta/\alpha)] \\ &\quad = 2I_n([(\beta - n/\alpha)^2 - k^2]^{1/2}a)K_n([(\beta - n/\alpha)^2 - k^2]^{1/2}\rho) \\ &\qquad \text{if } (\beta - n/\alpha)^2 > k^2, \\ &\quad = \pi i J_n([k^2 - (\beta - n/\alpha)^2]^{1/2}a)H_n^{(1)}([k^2 - (\beta - n/\alpha)^2]^{1/2}\rho) \\ &\qquad \text{if } (\beta - n/\alpha)^2 < k^2. \end{aligned}\right\} \tag{2.11}$$

Now it can be shown [4, corollary 2.2] that $I_n(x)K_n(x)$ is a monotonic decreasing function of x which assumes the value $1/|\,2n\,|$ at $x = 0$. Since, in addition, $K_n(x)$ is monotonic decreasing in x, it follows that

$$I_n(\Gamma_n a)K_n(\Gamma_n \rho) < 1/|\,2n\,|$$

for $\rho > a$. Combining this with the above result and the inequality (2.8), we obtain for fixed β and fixed $\rho > a$

$$\left| \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ikR}/R) \exp\,[in(\sigma - \zeta/\alpha)] \right| < \tilde{c}/|\,n\,| \tag{2.12}$$

for sufficiently large $|\,n\,|$. It follows that $\lim_{L\to\infty} I(L, r)$ exists for $r < 1$ and $\rho > a$; in fact

$$\lim_{L\to\infty} I(L, r) = \sum_{n=-\infty}^{\infty} r^{|n|} 4\pi I_n(\Gamma_n a)K_n(\Gamma_n \rho)e^{in\varphi}. \tag{2.13}$$

If for a particular n, Γ_n is imaginary, then the corresponding term in (2.13) is to be replaced by $2\pi^2 i r^{|n|} J_n(\gamma_n \cdot a)H_n^{(1)}(\gamma_n \rho)e^{in\zeta}$ where $\gamma_n = [k^2 - (\beta - n/\alpha)^2]^{1/2}$. We shall in the future write the series as above and leave to the reader the interpretation for imaginary Γ_n.

In order to evaluate $\lim_{r\to 1-} \lim_{L\to\infty} I(L, r)$, we refer to an earlier paper [4, corollary 2.1]. It was there shown that for $n \geq 0$

$$\frac{n}{v^2}\left[1 + \frac{v^2}{n^2}\right]^{1/2} < \frac{-K_n'(v)}{vK_n(v)} < \frac{n}{v^2}\left[1 + \frac{v^2}{n(n-1)}\right]^{1/2}. \tag{2.14}$$

Hence

$$\frac{K_n'(v)}{K_n(v)} < -\frac{n}{v}\left[1 + \frac{v^2}{n^2}\right]^{1/2} < -\frac{n}{v}$$

and

$$K_n(v) < K_n(v_0)(v_0/v)^n \qquad \text{for} \qquad (v > v_0).$$

This coupled with the above mentioned fact that

$$I_n(v)K_n(v) \leq 1/2n$$

gives

$$I_n(\Gamma_n a)K_n(\Gamma_n \rho) \leq (a/\rho)^n/2n \tag{2.15}$$

for $\rho > a$. Since $I_{-n}(\Gamma_n a)K_{-n}(\Gamma_n \rho) = I_n(\Gamma_n a)K_n(\Gamma_n \rho)$ the series (2.13) is majorized uniformly in $0 \leq r \leq 1$ by

$$\sum_{n=-\infty}^{\infty}{}' (a/\rho)^{|n|}/|\,2n\,|.$$

Thus for $\rho > a$

$$\lim_{r\to 1-} \lim_{L\to\infty} I(L, r) = \sum_{n=-\infty}^{\infty} 4\pi e^{in\varphi} I_n(\Gamma_n a)K_n(\Gamma_n \rho). \tag{2.16}$$

Finally we shall establish (b), namely, that the $\lim_{r\to 1^-} I(L, r)$ exists uniformly in L. To this end, we define

$$h_\epsilon(r, \sigma) = P(r, \sigma) \qquad \text{for} \quad P(r, \sigma) \leq \epsilon,$$

$$= \epsilon \qquad \text{for} \quad P(r, \sigma) > \epsilon,$$

$$a(r, \epsilon) = 1 - \int_0^{2\pi} d\sigma\, h_\epsilon(r, \sigma)$$

$$f_\epsilon(r, \sigma) = h_\epsilon(r, \sigma)/a(r, \epsilon)$$

$$g_\epsilon(r, \sigma) = [P(r, \sigma) - h_\epsilon(r, \sigma)]/a(r, \epsilon).$$

It is readily seen that

$$\int_0^{2\pi} g_\epsilon(r, \sigma)\, d\sigma = 1$$

$$0 < f_\epsilon(r, \sigma) \leq \epsilon/a(r, \epsilon) \leq \epsilon/(1 - 2\pi\epsilon)$$

and

$$a(r, \epsilon) \to 1 \qquad \text{as} \qquad r \to 1^-.$$

Now $f_\epsilon(r, \sigma)$ and $g_\epsilon(r, \sigma)$ are continuous functions of period 2π. Further $f_\epsilon(r, \sigma)$ is of total variation $2\epsilon/a(r, \epsilon)$ over any interval of length 2π. On the other hand $g_\epsilon(r, \sigma)$ is a needle-like function, vanishing outside of an interval of length $\delta(\epsilon, r)$ about the origin, being otherwise positive valued and enclosing a unit area. It follows that any integral of the form $\int_0^{2\pi} g_\epsilon(r, \sigma)w(\sigma)\, d\sigma$ differs from $w(0)$ by less than the oscillation of $w(\sigma)$ in the non-zero interval of $g_\epsilon(r, \sigma)$. It follows from known properties of $P(r, \sigma)$ that $\delta(\epsilon, r)$ approaches zero for each ϵ as $r \to 1^-$.

We write $I(L, r)$ as the sum of two integrals

$$I(L, r) = a(r, \epsilon) \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re) f_\epsilon(r, \sigma - \zeta/\alpha) \tag{2.17}$$

$$+ a(r, \epsilon) \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re) g_\epsilon(r, \sigma - \zeta/\alpha)$$

Expanding $f_\epsilon(r, \sigma)$ in a Fourier series, we obtain

$$f_\epsilon(r, \sigma) = \sum_{n=-\infty}^{\infty} b_n(r, \epsilon) e^{in\sigma}, \tag{2.18}$$

where

$$| b_n(r, \epsilon) | \leq \text{Var}\,(f_\epsilon)/| n | \leq 2\epsilon/[| n | a(r, \epsilon)].$$

Substituting this into the first of the integrals of (2.17) and interchanging the order of summation and integration (finite limits) gives

$$a(r, \epsilon) \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re) f_\epsilon(r, \sigma - \zeta/\alpha) \tag{2.19}$$

$$= \sum_{n=-\infty}^{\infty} a(r, \epsilon) b_n(r, \epsilon) \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re) \exp\,[in(\sigma - \zeta/\alpha)].$$

Finally if we combine this with the inequalities (2.12) and (2.18) we see that the series (2.19) is majorized by

$$\epsilon \sum{}' \tilde{c}/n^2 \leq \epsilon C_1$$

which converges to zero with ϵ uniformly in L for fixed β and fixed $\rho > a$.

The second integral in (2.17) approaches $\int_{-L}^{L} d\zeta e^{i\beta\zeta}(e^{ikR}/R)$ as $r \to 1^-$. Further, we note that the oscillation of $e^{i\beta\zeta}(e^{ik\Re}/\Re)$ in the $\delta(\epsilon, r)$ interval about $\sigma = \zeta/\alpha$ is bounded by

$$\left| \frac{\partial}{\partial\sigma} (e^{ik\Re}/\Re) \right| \delta(\epsilon, r) \leq C_2\delta(\epsilon, r)/\Re^2.$$

Hence

$$\left| \int_{-L}^{L} d\zeta \int_0^{2\pi} d\sigma\, e^{i\beta\zeta}(e^{ik\Re}/\Re)g_\epsilon(r, \sigma - \zeta/\alpha) - \int_{-L}^{L} d\zeta\, e^{i\beta\zeta}(e^{ikR}/R) \right|$$

$$\leq C_2\delta(\epsilon, r) \int_{-\infty}^{\infty} \Re^{-2}\, d\zeta = C_3\delta(\epsilon, r)$$

uniformly in L. We have finally that

$$| I(L, r)/a(r, \epsilon) - \lim_{r\to 1-} I(L, r) | \leq C_1\epsilon/(1 - 2\pi\epsilon) + C_3\delta(\epsilon, r).$$

Since ϵ is arbitrary and $\delta(\epsilon, r)$ approaches zero as $r \to 1^-$, it follows that $\lim_{r\to 1-} I(L, r)$ exists uniformly in L. This concludes the proof of (b). Combining this with (a) it follows from a well known theorem on interchange of limits that

$$I = \lim_{L\to\infty} \lim_{r\to 1-} I(L, r) = \lim_{r\to 1-} \lim_{L\to\infty} I(L, r)$$

$$= \sum_{n=-\infty}^{\infty} 4\pi e^{in\varphi} I_n(\Gamma_n a) K_n(\Gamma_n \rho) \tag{2.20}$$

for $\rho > a$.

It should be remarked that for a point inside of the helix, i.e., $\rho < a$, the vector potential components can be obtained by simply interchanging a and ρ since these two variables enter symmetrically in the integral I.

3. Electromagnetic field expressions. In the previous section it was shown that the integral I converges for $\rho \neq a$. We now define

$$I(\rho, \varphi) = \int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}(e^{ikR}/R), \tag{3.1}$$

where $R = [\zeta^2 + a^2 + \rho^2 - 2a\rho \cos(\varphi - \zeta/\alpha)]^{1/2}$ and β is real. It remains to show that this integral converges for $\rho = a$ provided that $\varphi \neq 0$ (that is for a point not on the helix itself). This will be a by-product of our investigation into the derivatives of $I(\rho, \varphi)$ with respect to ρ and φ.

For any pair of values (ρ, φ) different from $(a, 0)$ there is a neighborhood of this point in which R is bounded away from zero. Further R is essentially equal to $|\zeta|$ for large $|\zeta|$. Hence for the points (ρ, φ) of such neighborhood any integral majorized by

M/R^{μ}, $(\mu > 1)$, will converge uniformly. This is clearly the case for all derivatives of the integrand of $I(\rho, \varphi)$ with respect to ρ and φ. It follows that

$$\int_{\rho_0}^{\rho} d\rho \int_{-\infty}^{\infty} d\zeta \, \partial_\rho[e^{i\beta\zeta}(e^{ikR}/R)] = \int_{-\infty}^{\infty} d\zeta \, \{e^{i\beta\zeta}[(e^{ikR}/R) - (e^{ikR_0}/R_0)]\}, \tag{3.2}$$

where R_0 is obtained from R by replacing ρ by ρ_0.

Suppose now that $\rho_0 = a$ whereas $\rho \neq a$. Then since $I(\rho, \varphi)$ converges, the right-hand side of (3.2) can be written as the sum of two convergent integrals; hence $I(a, \varphi)$ converges for $\varphi \neq 0$. Whether $\rho_0 = a$ or not, the integrand of the outer integral on the left-hand side of (3.2) is continuous. Hence

$$\partial_\rho I(\rho, \varphi) = \partial_\rho[I(\rho, \varphi) - I(\rho_0, \varphi)] = \partial_\rho\left\{\int_{\rho_0}^{\rho} d\rho \int_{-\infty}^{\infty} d\zeta \, \partial_\rho[e^{i\beta\zeta}(e^{ikR}/R)]\right\}$$

$$= \int_{-\infty}^{\infty} d\zeta \, \partial_\rho[e^{i\beta\zeta}(e^{ikR}/R)].$$

It is clear that a similar argument applies to all derivatives of $I(\rho, \varphi)$ with respect to ρ and φ. Further we now know that I converges for all points off of the wire. This is clearly independent of the coordinate system used to express the point. It follows therefore by precisely the same argument as above that the derivatives for the components of Π with respect to x, y, and z can be obtained by differentiating under the integral sign.

In order to evaluate the derivatives of $I(\rho, \varphi)$ for $\rho \neq a$, we recall the inequality (2.15) which shows that the series expansion for $I(\rho, \varphi)$ [see (2.20)] differentiated termwise with respect to φ, is majorized by $\sum (\rho/a)^{|n|}$ for $\rho < a$ and by $\sum (a/\rho)^{|n|}$ for $\rho > a$. It follows that $\partial_\varphi I(\rho, \varphi)$ can be obtained by termwise differentiation of the series expansion.

A similar result holds for $\partial_\rho I(\rho, \varphi)$. Here we make use of the inequality (2.14) for the case $\rho > a$. Thus

$$| K_n'(v) | \leq K_n(v) \frac{|n|}{v}\left[1 + \frac{v^2}{n(n-1)}\right]^{1/2}.$$

For $v \geq \Gamma_n a$, the functions $(|n|/v)[1 + v^2/n(n-1)]^{1/2}$ are uniformly bounded by some positive number M. Hence

$$| \Gamma_n I_n(\Gamma_n a) K_n'(\Gamma_n \rho) | \leq \Gamma_n M I_n(\Gamma_n a) K_n(\Gamma_n \rho) \leq M_1(a/\rho)^{|n|},$$

so that the series expansion for $I(\rho, \varphi)$ can be differentiated termwise with respect to ρ in case $\rho > a$.

Finally $\rho < a$, we make use of the corresponding inequality for the I_n function $(n \geq 0)$ (see [4, corollary 1.1])

$$\frac{n}{v^2}\left[1 + \frac{v^2}{n(n+1)}\right]^{1/2} \leq \frac{1}{v}\frac{I_n'(v)}{I_n(v)} \leq \frac{n}{v^2}\left[1 + \frac{v^2}{n^2}\right]^{1/2}. \tag{3.3}$$

For $v \leq \Gamma_n a$, the functions $(n/v)[1 + (v/n)^2]^{1/2}$ are uniformly bounded by M'/v for some positive number M'. Hence

$$| \Gamma_n I_n'(\Gamma_n \rho) K_n(\Gamma_n a) | \leq (M_1'/\rho)(\rho/a)^{|n|},$$

so that again the series expansion can be differentiated termwise with respect to ρ in the case $\rho < a$.

It can now be readily shown that the components of the vector potential (and hence the electromagnetic field components) satisfy the wave equation

$$\Delta\Phi - \epsilon\mu\partial_{tt}\Phi = 0$$

at all points (x, y, z) not on the helix. In fact, since we can differentiate under the integral sign this amounts simply to showing that the quantity $e^{-i\omega t}(e^{ikr}/r)$ where $r = [(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2]^{1/2}$ satisfies the wave equation (Note that $k^2 = \epsilon\mu\omega^2$).

We shall conclude this section by obtaining explicit expressions for the field components from the formulae (1.2) and (1.6). We first transform the integrals of (1.6) as in (1.7). Then we have

$$\begin{aligned}
\Pi_\rho &= aFe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}(e^{ikr}/r)\sin(\varphi - z/\alpha - \zeta/\alpha),\\
\Pi_\varphi &= aFe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}(e^{ikr}/r)\cos(\varphi - z/\alpha - \zeta/\alpha), \qquad (3.4)\\
\Pi_z &= \alpha Fe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}(e^{ikr}/r),
\end{aligned}$$

where

$$r = [\zeta^2 + a^2 + \rho^2 - 2a\rho\cos(\varphi - z/\alpha - \zeta/\alpha)]^{1/2}.$$

We shall compute E explicitly. Now

$$\nabla\cdot\Pi = \frac{1}{\rho}\Pi_\rho + \partial_\rho\Pi_\rho + \frac{1}{\rho}\partial_\varphi\Pi_\varphi + \partial_z\Pi_z$$

$$\partial_\rho\Pi_\rho = aFe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}r^{-1}[\rho - a\cos(\varphi - z/\alpha - \zeta/\alpha)]\sin(\varphi - z/\alpha - \zeta/\alpha)\partial_r(e^{ikr}/r),$$

$$\begin{aligned}
\partial_\varphi\Pi_\varphi = aFe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}r^{-1}[&a\rho\sin(\varphi - z/\alpha - \zeta/\alpha)\cos(\varphi - z/\alpha - \zeta/\alpha)\partial_r(e^{ikr}/r)\\
&- \sin(\varphi - z/\alpha - \zeta/\alpha)(e^{ikr}/r)],
\end{aligned}$$

$$\partial_z\Pi_z = i\beta\Pi_z + \alpha Fe^{i\beta z}\int_{-\infty}^{\infty} d\zeta\, e^{i\beta\zeta}(\alpha r)^{-1}[-a\rho\sin(\varphi - z/\alpha - \zeta/\alpha)]\partial_r(e^{ikr}/r).$$

Combining the terms in $\nabla\cdot\Pi$ and making the appropriate cancellations gives

$$\nabla\Pi = i\beta\Pi_z \qquad (3.5)$$

Since $E = \nabla(\nabla\cdot\Pi) + k^2\Pi$, we have for $\rho > a$

$$E_\rho = i\beta\partial_\rho \Pi_z + k^2\Pi_\rho$$

$$= 4\pi i\beta\alpha F e^{i\beta z} \sum_{n=-\infty}^{\infty} \Gamma_n I_n(\Gamma_n a)K_n(\Gamma_n \rho) \exp[in(\varphi - z/\alpha)]$$

$$- i2\pi k^2 a F e^{i\beta z} \sum_{n=-\infty}^{\infty} [I_{n-1}(\Gamma_n a)K_{n-1}(\Gamma_n \rho) \tag{3.6a}$$

$$- I_{n+1}(\Gamma_n a)K_{n+1}(\Gamma_n \rho)] \exp[in(\varphi - z/\alpha)],$$

$$E_\varphi = \frac{i\beta}{\rho}\partial_\varphi \Pi_z + k^2\Pi_\varphi$$

$$= 4\pi i\beta\alpha F \rho^{-1} e^{i\beta z} \sum_{n=-\infty}^{\infty} inI_n(\Gamma_n a)K_n(\Gamma_n \rho) \exp[in(\varphi - z/\alpha)] \tag{3.6b}$$

$$+ 2\pi k^2 a F e^{i\beta z} \sum_{n=-\infty}^{\infty} [I_{n-1}(\Gamma_n a)K_{n-1}(\Gamma_n \rho)$$

$$+ I_{n+1}(\Gamma_n a)K_{n+1}(\Gamma_n \rho)] \exp[in(\varphi - z/\alpha)],$$

$$E_z = i\beta\partial_z \Pi_z + k^2\Pi_z$$

$$= 4\pi(k^2 - \beta^2)\alpha F e^{i\beta z} \sum_{n=-\infty}^{\infty} I_n(\Gamma_n a)K_n(\Gamma_n \rho) \exp[in(\varphi - z/\alpha)] \tag{3.6c}$$

$$- 4\pi i\beta F e^{i\beta z} \sum_{n=-\infty}^{\infty} in\, I_n(\Gamma_n a)K_n(\Gamma_n \rho) \exp[in(\varphi - z/\alpha)].$$

The electric field components for $\rho < a$ are very similar. One can obtain the magnetic field components directly from

$$H = -i\omega\epsilon \nabla \times \Pi.$$

Since we shall have no occasion to use these in the present paper, we refrain from writing them out.

4. Determination of the propagation constant. In this section we shall determine the propagation constant β so that the electric field along the helix is "essentially" zero. Since we assume the helix to be a perfect conductor this is the proper boundary condition for our problem. The electric field in the helical direction at any point (ρ, φ, z) is

$$E_h = (\alpha/[a^2 + \alpha^2]^{1/2})E_z + (a/[a^2 + \alpha^2]^{1/2})E_\varphi . \tag{4.1}$$

In order to approach the helix radially, we set $\varphi = z/a$ in (4.1) and try to determine β so that $\lim_{\rho\to a} E_h = 0$. We will suppose that $\rho > a$; the case $\rho < a$ is entirely similar.

Because of the singular nature of the field about an infinitely thin conductor we cannot actually choose β so that $\lim_{\rho\to a} E_h = 0$. Since E becomes infinite as $\rho \to a$, the best we can do is to determine the proper phase velocity, β_0, so that

$$\lim_{\rho\to a} \frac{E_h(\beta_0)}{E_h(|\beta| \neq |\beta_0|)} = 0. \tag{4.2}$$

For $\varphi = z/\alpha$ and $\rho > a$, E_h can be expressed as

$$[a^2 + \alpha^2]^{1/2}E_h = 4\pi\alpha^2(k^2 - \beta^2)Fe^{i\beta z}\sum_{n=-\infty}^{\infty} I_n(\Gamma_n a)K_n(\Gamma_n\rho)$$

$$+ 2\pi a^2k^2Fe^{i\beta z}\sum_{n=-\infty}^{\infty}[I_{n-1}(\Gamma_n a)K_{n-1}(\Gamma_n\rho) + I_{n+1}(\Gamma_n a)K_{n+1}(\Gamma_n\rho)] \quad (4.3)$$

$$-4\pi i\beta\alpha(1 - a/\rho)Fe^{i\beta z}\sum_{n=-\infty}^{\infty} in\, I_n(\Gamma_n a)K_n(\Gamma_n\rho).$$

We shall first prove that the last term in (4.3) approaches zero as $\rho \to a$. In its original integral form this term becomes

$$J(\rho) = (1 - a/\rho)(a\rho)^{-1}\partial_\varphi[I(\rho, \varphi)]_{\varphi=0} = (1 - a/\rho)\int_{-\infty}^{\infty} e^{i\beta\zeta}R^{-1}\sin\zeta\,\partial_R(e^{ikR}/R)\,d\zeta$$

aside from a factor independent of ρ. Here

$$R = [\zeta^2 + a^2 + \rho^2 - 2a\rho\cos\zeta/\alpha]^{1/2}$$

is an even function of ζ. The $e^{i\beta\zeta}$ can be expanded as

$$e^{i\beta\zeta} = \cos\beta\zeta + i\sin\beta\zeta.$$

Since $[\cos\beta\zeta R^{-1}\sin\zeta\,\partial_R(e^{ikR}/R)]$ is odd, its integral vanishes.
Hence

$$J(\rho) = (1 - a/\rho)\int_{-\infty}^{\infty} R^{-1}\sin\beta\zeta\sin\zeta\,\partial_R(e^{ikR}/R)\,d\zeta.$$

We note that

$$|\zeta| \leq [\zeta^2 + (a - \rho)^2]^{1/2} \leq R.$$

It follows that there exist constants c_1 and c_2 such that

$$|\partial_R(e^{ikR}/R)| \leq c_1/|\zeta| \qquad \text{for} \quad |\zeta| > 1,$$

$$\leq c_2/[\zeta^2 + (a - \rho)^2] \qquad \text{for} \quad |\zeta| \leq 1 \quad \text{and} \quad |\rho - a| < 1.$$

Therefore we have

$$|J(\rho)| \leq (1 - a/\rho)\left\{2\int_1^{\infty} d\zeta\, c_1/\zeta^2 + 2\int_0^1 d\zeta\,\beta c_2\zeta^2[\zeta^2 + (a - \rho)^2]^{-3/2}\right\}$$

$$= (1 - a/\rho)\{2c_1 + 2\beta c_2[\log\{(1 + [1 + (a - \rho)^2]^{1/2})/(\rho - a)\}$$

$$- \{1 + (a - \rho)^2\}^{-1/2}]\},$$

which approaches zero as $\rho \to a^+$.

In determining the limit of E_h as $\rho \to a^+$, we can therefore omit the last term in (4.3). We denote by E_h' the thus abbreviated E_h. We now define the function

$$\Omega(\beta, \rho) = \sum_{n=-\infty}^{\infty} I_n(\Gamma_n a)K_n(\Gamma_n\rho). \quad (4.4)$$

Then for $\varphi = z/\alpha$ and $\rho > a$, we have

$$E_h' = \frac{4\pi F}{[a^2 + \alpha^2]^{1/2}} e^{i\beta z}[\alpha^2(k^2 - \beta^2)\Omega(\beta, \rho) + (a^2k^2/2)\{\Omega(\beta - 1/\alpha, \rho) + \Omega(\beta + 1/\alpha, \rho)\}]. \tag{4.5}$$

It turns out that

$$\lim_{\rho \to a^+} \Omega(\beta, \rho) = \infty, \tag{4.6}$$

so that we can hardly expect to obtain a β for which $\lim_{\rho \to a^+} E_h = 0$. However we shall also prove that

$$\lim_{\rho \to a^+} \Omega(\beta_1, \rho)/\Omega(\beta_2, \rho) = 1. \tag{4.7}$$

Hence

$$\lim_{\rho \to a^+} \frac{E_h(\beta)}{E_h(\beta_1)} = \frac{\alpha^2(k^2 - \beta^2) + a^2k^2}{\alpha^2(k^2 - \beta_1^2) + a^2k^2};$$

and for $\beta_1^2 \neq \beta^2$. This vanishes if and only if

$$\alpha^2(k^2 - \beta^2) + a^2k^2 = 0.$$

In other words the correct propagation constant β_0 is simply

$$\beta_0 = \pm k[a^2 + \alpha^2]^{1/2}\alpha^{-1}. \tag{4.8}$$

We now prove (4.6). In order to do this we make use of the inequalities (2.14) and (3.3), from which it follows that

$$-\frac{K_n'(v)}{vK_n(v)} - \frac{I_n'(v)}{vI_n(v)} < \frac{n}{v^2}\left\{\left[1 + \frac{v^2}{n(n-1)}\right]^{1/2} - \left[1 + \frac{v^2}{n(n+1)}\right]^{1/2}\right\}.$$

The expression on the right is monotonic decreasing achieving its maximum of $1/(n^2 - 1)$ at $v = 0$. Hence

$$\frac{(I_n(v)K_n(v))'}{I_n(v)K_n(v)} > -v/(n^2 - 1)$$

and since $I_n(0)K_n(0) = 1/(2n)$, we obtain

$$(2n)^{-1} \geq I_n(v)K_n(v) \geq (2n)^{-1} \exp\{-v^2/[2(n^2 - 1)]\}.$$

Finally for sufficiently large $|\, n \,|$,

$$\Gamma_n = [(\beta - n/\alpha)^2 - k^2]^{1/2} > c\,|\, n \,|$$

and hence

$$I_n(\Gamma_n a)K_n(\Gamma_n a) \geq |\, 2n \,|^{-1} \exp(-c^2).$$

Since $\Omega(\beta, \rho)$ consists of positive continuous terms, monotonically increasing as $\rho \to a^+$, we have

$$\lim_{\rho \to a^+} \Omega(\beta, \rho) = \Omega(\beta, a) \gg (1/2) \exp(-c^2) \sum_{n=-\infty}^{\infty}{}' |\, n \,|^{-1},$$

which diverges.

We shall next establish (4.7) also by means of the inequalities (2.14) and (3.3). Now clearly

$$\frac{\partial_v[I_n(va)K_n(v\rho)]}{I_n(va)K_n(v\rho)} = a\frac{I_n'(va)}{I_n(va)} + \rho\frac{K_n'(v\rho)}{K_n(v\rho)}$$

$$= v\left\{a^2\left[\frac{I_n'(va)}{vaI_n(va)}\right] - \rho^2\left[\frac{-K_n'(v)}{vK_n(v)}\right]\right\}.$$

It follows from (2.14) and (3.3) that for $n \geq 1$ and $v > 0$

$$\frac{\partial_v[I_n(va)K_n(v\rho)]}{I_n(va)K_n(v\rho)} \leq v\left\{a^2\frac{n}{(va)^2}\left[1 + \frac{(va)^2}{n^2}\right]^{1/2} - \rho^2\frac{n}{(v\rho)^2}\left[1 + \frac{(v\rho)^2}{n^2}\right]^{1/2}\right\}$$

$$\leq \frac{n}{v}\left\{\left[1 + \frac{(va)^2}{n^2}\right]^{1/2} - \left[1 + \frac{(v\rho)^2}{n^2}\right]^{1/2}\right\} < 0$$

for $\rho > a$. Likewise ($n \geq 1$ and $v \geq 0$)

$$\frac{\partial_v[I_n(va)K_n(v\rho)]}{I_n(va)K_n(v\rho)} \geq v\left\{\frac{n}{v^2}\left[1 + \frac{(va)^2}{n(n+1)}\right]^{1/2} - \frac{n}{v^2}\left[1 + \frac{(v\rho)^2}{n(n-1)}\right]^{1/2}\right\}.$$

It is easily shown that the function within the brackets is negative monotonic increasing for $\rho > a$, achieving a minimum of

$$\left(-\frac{n(\rho^2 - a^2) + (\rho^2 + a^2)}{n^2 - 1}\right) \quad \text{at} \quad v = 0.$$

Hence

$$\frac{\partial_v[I_n(va)K_n(v\rho)]}{I_n(va)K_n(v\rho)} \geq -\omega_n v$$

where

$$\omega_n = \frac{n(\rho^2 - a^2) + (\rho^2 + a^2)}{n^2 - 1}.$$

If we now integrate the above logarithmic derivative from $v > 0$ to $v + \Delta > 0$ we obtain for $n \geq 1$

$$\exp\left(-\frac{\omega_n}{2}[(v+\Delta)^2 - v^2]\right) \leq \frac{I_n[(v+\Delta)a]K_n[(v+\Delta)\rho]}{I_n(va)K_n(v\rho)} \leq 1 \quad \text{for} \quad \Delta \geq 0 \quad (4.9a)$$

and

$$1 \leq \frac{I_n[(v+\Delta)a]K_n[(v+\Delta)\rho]}{I_n(va)K_n(v\rho)} \leq \exp\left(-\frac{\omega_n}{2}[(v+\Delta)^2 - v^2]\right) \quad \text{for} \quad \Delta \leq 0. \quad (4.9b)$$

Finally because $I_{-n}(va)K_{-n}(v\rho) = I_n(va)K_n(v\rho)$, the above inequalities hold for $n \leq -1$ if we replace n in ω_n by $|n|$. We can combine (4.9a) and (4.9b) into

$$e^{-\epsilon(\rho,n)} \leq \frac{I_n[(v+\Delta)a]K_n[(v+\Delta)\rho]}{I_n(va)K_n(v\rho)} \leq e^{\epsilon(\rho,n)} \quad (4.10)$$

for all $|n| > 1$, $v > 0$, $(v + \Delta) > 0$, and $\rho > a$, where

$$\epsilon(\rho, n) = \frac{|n|(\rho^2 - a^2) + \rho^2 + a^2}{n^2 - 1} |(v + \Delta)^2 - v^2|.$$

It will be seen from the expression for $\Omega(\beta, \rho)$ that the appropriate values for v and $(v + \Delta)$ in $I_n(va)K_n(v\rho)$ are

$$v_n = \left[\left(\beta_1 - \frac{n}{\alpha}\right)^2 - k^2\right]^{1/2} = \Gamma_n^{(1)}$$

and

$$(v + \Delta)_n = \left[\left(\beta_2 - \frac{n}{\alpha}\right)^2 - k^2\right]^{1/2} = \Gamma_n^{(2)}.$$

Hence

$$(v + \Delta)_n^2 - v_n^2 = \beta_2^2 - \beta_1^2 - 2n(\beta_2 - \beta_1)/\alpha.$$

It is clear that $\epsilon(\rho, n)$ is the sum of two parts one of which is of the order of $(\rho - a)$ and the other of the order of $1/(|n|)$. Hence given β_1, β_2, and $\eta > 0$, there exists an N and a $\delta > 0$ such that for $|n| \geq N$ and $0 < \rho - a < \delta$

$$\epsilon(\rho, n) < \eta.$$

The finite sums $\sum_{-N}^{N} I_n(\Gamma_n^{(i)}a)K_n(\Gamma_n^{(i)}\rho)$, remain uniformly bounded as $\rho \to a^+$. Hence

$$1 - \frac{\Omega(\beta_1, \rho)}{\Omega(\beta_2, \rho)} = \frac{\sum_{-N}^{N}\{\ \} + \sum_{-\infty}^{-N}\{\ \} + \sum_{N}^{\infty}\{\ \}}{\Omega(\beta_2, \rho)}$$

$$= \frac{\sum_{-N}^{N}\{\ \}}{\Omega(\beta_2, \rho)} + \frac{\sum_{-\infty}^{-N}\{\ \} + \sum_{N}^{\infty}\{\ \}}{\Omega(\beta_2, \rho)},$$

where $\{\ \}$ stands for the running term

$$\{I_n(\Gamma_n^{(2)}a)K_n(\Gamma_n^{(2)}\rho) - I_n(\Gamma_n^{(1)}a)K_n(\Gamma_n^{(1)}\rho)\};$$

thus

$$\left|1 - \frac{\Omega(\beta_1, \rho)}{\Omega(\beta_2, \rho)}\right| \leq \frac{\left|\sum_{-N}^{N}\{\ \}\right|}{\Omega_2(\beta, \rho)} + \frac{(e^{\eta} - e^{-\eta})\left[\Omega(\beta_2, \rho) - \sum_{-N}^{N} I_n(\Gamma_n^{(2)}a)K_n(\Gamma_n^{(2)}\rho)\right]}{\Omega(\beta_2, \rho)}.$$

It follows that

$$\overline{\lim_{\rho \to a^+}}\left|1 - \frac{\Omega(\beta_1, \rho)}{\Omega(\beta_2, \rho)}\right| \leq 0 + e^{\eta} - e^{-\eta},$$

and since η is arbitrarily small, we have

$$\lim_{\rho \to a^+} \frac{\Omega(\beta_1, \rho)}{\Omega(\beta_2, \rho)} = 1.$$

5. **Conclusion.** The net result of our investigation has been to obtain expressions for the electromagnetic field which satisfy the Maxwell equations away from the helix and

which satisfy the appropriate boundary conditions along the helix. The propagation constant in the axial direction was found (4.8) to be

$$\beta_0 = \pm k[a^2 + \alpha^2]^{1/2}\alpha^{-1} \tag{5.1}$$

which corresponds to precisely the phase velocity one would obtain if the wave traveled along the helix itself with the free space velocity of propagation.
In case

$$(\beta_0 - n/\alpha)^2 > k^2 \tag{5.2}$$

for all n ($n = 0, \pm 1, \pm 2, \cdots$) none of the terms in the expansion of the field equations are unmodified Hankel functions. As a consequence there is no net radial flow of energy. On the other hand for a given helix there is a series of frequency bands for which the inequality (5.2) is not satisfied for all n; in this case a finite number of Hankel function terms do appear in the field expressions. It can be shown that these terms do cause a net radial flow of energy. One might expect this phenomena to be reflected in the propagation constant β_0 ; that is, one would expect β_0 to be complex. More precisely, unless there is a source of energy within the wire itself, one would expect

$$Im(\beta_0) = \frac{1}{2}\frac{\text{Energy radiated per unit distance in the } z\text{-direction}}{\text{Energy flow in the } z\text{-direction}} . \tag{5.3}$$

However the computed value of β_0 (5.1) is actually consistent with (5.3) for it can be shown that whereas the energy radiated per unit distance in the z-direction is finite, the energy flow in the z-direction is infinite. As a matter of fact not only is the total contribution of the set of terms satisfying (5.2) infinite, but the contribution of each of the radiating terms to the energy flow in the z-direction is itself infinite.

In spite of the physical reasonableness of our result, there is, in the case of the radiating helix, reason to be dissatisfied with it. If (5.2) is satisfied for all n, the expansion which we have obtained can be used for helices with conductors of finite thickness. One could, for instance, replace the delta function of section 2 by a finite needle function (say $P(r_0 , \sigma)$ with fixed $r_0 < 1$). The helix would then be constructed from infinitely thin tape. For instance Π_z would be of the type ($\rho > a$)

$$\Pi_z = \alpha F e^{i\beta z} \sum_{n=-\infty}^{\infty} a_n I_n(\Gamma_n a) K_n(\Gamma_n \rho) \exp\,[in(\varphi - z/\alpha)]$$

where the (a_n) are Fourier coefficients of the needle function. Repeating the above development one would obtain a real β_0 for the a_n's sufficiently close to 1. However, in the radiating case, one would not be able to obtain a solution in this way. In this case, no real β would allow the boundary conditions to be satisfied whereas the entire development (starting with the vector potential integrals) is meaningless for complex β. This is also reflected by the fact that a complex β which satisfies the boundary conditions would make the arguments of all of the Hankel functions complex so that the field would decay exponentially. As a consequence we would have no radiating terms which would be contrary to (5.3). Thus in case (5.2) is not satisfied by all n, one would expect a real helical wire to radiate. However one would not expect the field expressions for finite conductors to have expansions of the type which we have obtained for this case.

References

1. Franz Ollendorff, *Die Grundlagen der Hochfrequenztechnik*, Springer, Berlin, 1926, pp. 79-87.

2. J. R. Pierce, *Theory of the beam-type traveling-wave tube*, Proceedings of the Institute of Radio Engineers **35**, 111-123 (1947).

3. R. S. Phillips and Henry Malin, *A helical wave guide*, New York University Mathematical Research Group, Report No. 170-3, Army Air Forces, Watson Laboratories, 1947.

4. R. S. Phillips and Henry Malin, *Bessel function approximations*, to be published in the Amer. J. Math.

5. S. Ramo and J. R. Whinnery, *Fields and waves in modern radio*, Wiley and Sons, New York, 1944.

6. G. N. Watson, *Theory of Bessel functions*, MacMillan, New York, 1944.

7. A. Zygmund, *Trigonometrical series*, Monografje Matematyczne, Warsaw, 1935.

TRANSIENTS IN MULTIPLY PERIODIC NON-LINEAR SYSTEMS*

BY

FRANK E. BOTHWELL

The University of Chicago

1. Introduction. The transient behavior of the general dynamical system of n degrees of freedom, subject to certain restrictions, will be treated. The system must be (a) nearly linear, (b) nearly conservative, (c) nearly autonomous; the deviation from the case of complete linearity, conservativeness, and autonomy being indicated by the magnitude of a parameter μ. It is further assumed that the explicit dependence of the non-autonomous terms on time is in the form of sinusoids of frequencies Ω_1 , Ω_2 , $\cdots$, Ω_m .

The general system is transformed to normal coordinates and thence to elliptical coordinates, the resulting form being convenient for the development of the solution for small values of μ. An auxiliary set of differential equations is so constructed that its singular points determine the limit cycles of the original system for the limiting case, $\mu \to 0$. The stability of each limit cycle in the original system is the same as the stability of the corresponding singular point in the auxiliary system. Furthermore, the coordinates of the auxiliary set are the amplitudes of the n fundamental frequency components of the original system, so that the motion of the representative point in the phase space of the auxiliary system describes approximately the transient behavior of the amplitudes of the fundamental frequency components of the original system for small values of the parameter μ.

The general cases of internal and external resonances, in which two or more fundamental frequency components are commensurate, are treated. Finally, the method is applied to a problem of practical interest.

2. The method. Under restrictions (a), (b), and (c), the equations which describe the general system are

$$\sum_{j=1}^{n} \left(a_{ij} \frac{d^2 y_j}{dt^2} - b_{ij} y_j \right) + \mu g_i = 0, \qquad i = 1, 2, \cdots, n. \tag{1}$$

The quantities g_i are functions of the parameter μ, the independent variable t, and the n dependent variables y_j and their various derivatives. The final restriction on the system (1) is that there exist a closed region D, including the origin, in which each g_i is bounded for all values of t and for sufficiently small values of μ.

By means of a suitable linear transformation [1]

$$y_j = \sum_{k=1}^{n} c_{jk} x_k , \qquad j = 1, 2, \cdots, n,$$

Eqs. (1) may be transformed to normal coordinates. Thus

$$\frac{d^2 x_i}{dt^2} + \omega_i^2 x_i + \mu f_i = 0, \qquad i = 1, 2, \cdots, n. \tag{2}$$

The constants ω_i^2 are the roots of the characteristic equation of the linear system which

*Received July 22, 1949.

occurs when μ is set equal to zero in (1). Since the kinetic energy of the linear system is a positive definite form, the constants ω_i^2 are positive [1].

The classical approach to equations (2) involves the expansion of the normal modes x_i and their fundamental frequencies in power series in μ. This procedure is an extension of the method presented by Kryloff and Bogoliuboff for a system of a single degree of freedom [2]. Concerning this extension, only one remark will be made. In the system of a single degree of freedom, secular terms in the series solution may be avoided by satisfying two equations at each stage of the process. One of the two equations determines the amplitude of the corresponding term in the series, and the other equation determines the contribution of the corresponding term to the fundamental frequency of oscillation. In the case of a multiply periodic system, each single equation is replaced by a set of n simultaneous non-linear equations.

For the purpose of this note, it is convenient to use elliptical coordinates defined by the equations

$$\begin{aligned} x_i &= r_i \cos \theta_i \\ \frac{dx_i}{dt} &= \omega_i r_i \sin \theta_i \end{aligned} \qquad i = 1, 2, \cdots, n. \tag{3}$$

In these coordinates, Eqs. (2) become

$$\begin{aligned} \frac{dr_i}{dt} &= -\frac{\mu}{\omega_i} f_i \sin \theta_i \\ \frac{d\theta_i}{dt} &= -\omega_i - \frac{\mu}{\omega_i r_i} f_i \cos \theta_i \end{aligned} \qquad i = 1, 2, \cdots, n. \tag{4}$$

By means of the change of variable,

$$T = \mu t,$$

Eqs. (4) may be written

$$\begin{aligned} \frac{dr_i}{dT} &= -\frac{1}{\omega_i} f_i \sin \theta_i \\ \frac{d\theta_i}{dT} &= -\frac{\omega_i}{\mu} - \frac{1}{\omega_i r_i} f_i \cos \theta_i \end{aligned} \qquad i = 1, 2, \cdots, n. \tag{5}$$

Inasmuch as each g_i is bounded in D, each f_i must be bounded in D. Since the quantities f_i are functions of the variables x_i and their derivatives, and since the latter are periodic functions of the variables θ_i , the right sides of (5) are periodic functions of each variable θ_i . Furthermore, by hypothesis, each f_i is a periodic function of each variable ϕ_i where

$$\phi_i = -\frac{\Omega_i}{\mu} T + \phi_{i0}, \qquad i = 1, 2, \cdots, m.$$

As the parameter μ is decreased, the right sides of (5) take the form of certain slowly varying functions of T(due to the variation of r_i with time) modulated by rapidly varying quasi-periodic functions of T(due to the variations of θ_i and ϕ_i with time). Furthermore, the quasi-periods of the modulation approach zero with μ.

Let $K_i(r_1, r_2, \cdots r_n, \mu)$ represent the unmodulated portion of $f_i \sin \theta_i$, defined by

$$K_i(\mu) = \lim_{P\to\infty} \frac{1}{P} \int_0^P f_i \sin \theta_i \, dt, \qquad i = 1, 2, \cdots, n, \tag{6}$$

in which the variables r_i are considered constant in the integration. When evaluating the integral (6) for the limiting case $\mu \to 0$, the integrand is a function of each r_i (considered constant in the integration), a sinusoidal function of each θ_i (equal to $-\omega_i t + \theta_{i0}$ in the limiting case), and a sinusoidal function of each ϕ_i (equal to $-\Omega_i t + \phi_{i0}$).

Now the quantities

$$G_i = f_i \sin \theta_i - K_i(\mu), \qquad i = 1, 2, \cdots, n,$$

are quasi-periodic functions of T, the quasi-periods of which approach zero with μ. Furthermore, each G_i is bounded in D, so that the limit as $\mu \to 0$ of the indefinite integral of G_i is identically zero. Therefore, for the limiting case, $\mu \to 0$, Eqs. (5) may be written

$$\frac{dr_i}{dT} = -\frac{K_{i(0)}}{\omega_i}, \qquad i = 1, 2, \cdots, n,$$

and for small values of μ, (4) may be approximated by

$$\frac{dr_i}{dt} = -\frac{\mu}{\omega_i} K_{i(0)}, \qquad i = 1, 2, \cdots, n. \tag{7}$$

The system of n simultaneous non-linear differential equations (7) comprises the set which is auxiliary to the system (4). In fact, the variables of (7) are the amplitudes of the set (4). Therefore, the singular points of the former, defined by the equations

$$K_i(r_{10}, r_{20}, \cdots, r_{n0}, 0) = 0, \qquad i = 1, 2, \cdots, n, \tag{8}$$

fix the amplitudes of oscillation on the limit cycles of (4). Furthermore, the stability of each singular point of (7) is the same as the stability of the corresponding limit cycle of (4).

Let us consider first the non-resonant case, that is, the case in which no two of the set of frequencies ω_i and Ω_i are commensurate. In this case, (6) may be replaced by the expressions

$$K_i(0) = \frac{1}{(2\pi)^{n+m}} \int_0^{2\pi} \int_0^{2\pi} \cdots \int_0^{2\pi} f_i \sin \theta_i \, d\theta_1 \, d\theta_2 \cdots d\theta_n \, d\phi_1 \, d\phi_2 \cdots d\phi_m, \tag{9}$$

$$i = 1, 2, \cdots, n.$$

Thus the values of K_i, the singular points of the auxiliary system (7), and hence the limit cycles of (4) for small μ are independent of the phases θ_{i0} and ϕ_{i0} of the autoperiodic and heteroperiodic oscillations.

3. Internal and external resonances. If two frequences, say ω_1 and ω_2 are commensurate, that is, if there exist two positive integers n_1 and n_2 such that

$$n_1\omega_1 = n_2\omega_2,$$

the system will be said to have a simple resonance of order n_1/n_2 or n_2/n_1, whichever is smaller. On the other hand, if there exist k positive integers, $n_1, n_2, \cdots, n_k$, such that

$$n_1\omega_1 = n_2\omega_2 = \cdots = n_k\omega_k, \tag{10}$$

the system will be said to have a complex resonance of degree k, there being associated with the resonance $k - 1$ orders, n_2/n_1 , n_3/n_1 , $\cdots$, n_k/n_1 , where n_1 is the largest of the integers.

In this case, the integral (6) may be replaced by

$$K_i(0) = \frac{1}{2\pi^{n+m-k+1}} \int_0^{2\pi} \int_0^{2\pi} \cdots \int_0^{2\pi} f_i \sin \theta_i \, dx \, d\theta_{k+1} \cdots d\theta_n \, d\phi_1 \cdots d\phi_m , \qquad (11)$$

$$i = 1, 2, \cdots, n,$$

in which

$$\theta_i = \frac{N}{n_i} x + \theta_{i0} , \qquad i = 1, 2, \cdots, k,$$

where

$$N = \prod_{j=1}^{k} n_j .$$

Thus the integration on x from zero to 2π insures that θ_1 , θ_2 , $\cdots$ θ_k range over integral numbers of their respective periods.

If only frequencies ω_i and not Ω_i are involved in (10), the resonance is said to be internal. If only frequencies Ω_i and not ω_i are involved, the resonance will be called external. If both ω_i and Ω_i are involved, the resonance will be called compound.

In the general case, one may have a system with several complex compound resonances. In any case, it is always possible to find a set of integers (sufficiently large) such that (10) is approximately true. Therefore, the resonance case makes sense only provided the integers are all small. The values of the integrals (11) in many cases depend upon the phases θ_{i0} and ϕ_{i0} .

4. Example. Consider the circuit of Fig. 1, containing the non-linear resistor

$$R = -\mu(1 - \gamma^2 I^2). \qquad (12)$$

The resistance R is composed of a passive resistor in series with the negative resistance of a vacuum tube. The passive resistor is so chosen that the total resistance of the circuit at zero current (that is, $-\mu$) is small and negative. The term $\gamma^2 I^2$ is due to saturation of the tube.

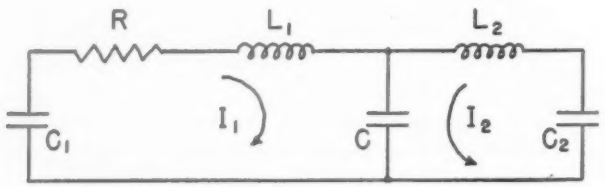

FIG. 1.

The two degrees of freedom of the circuit of Fig. 1 may be uncoupled by the transformation to normal coordinates

$$I_1 = I_1' + I_2'$$

$$I_2 = \alpha I_1' + \beta I_2' ,$$

in which

$$\alpha = \frac{1}{2\omega_{12}^2}\left[\omega_{22}^2 - \omega_{11}^2 + \{(\omega_{22}^2 - \omega_{11}^2)^2 + 4\omega_{12}^2\omega_{21}^2\}^{1/2}\right],$$

$$\beta = \frac{1}{2\omega_{12}^2}\left[\omega_{22}^2 - \omega_{11}^2 - \{(\omega_{22}^2 - \omega_{11}^2)^2 + 4\omega_{12}^2\omega_{21}^2\}^{1/2}\right],$$

where

$$\omega_{11}^2 = \frac{1}{L_1}\left(\frac{1}{c_1} + \frac{1}{c}\right), \qquad \omega_{12}^2 = \frac{1}{L_1 c},$$

$$\omega_{22}^2 = \frac{1}{L_2}\left(\frac{1}{c_2} + \frac{1}{c}\right), \qquad \omega_{21}^2 = \frac{1}{L_2 c}.$$

The resulting circuit is shown in Fig. 2.

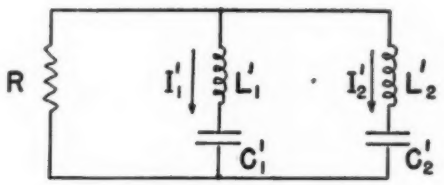

FIG. 2.

The equivalent circuit elements are given by the expressions

$$L_1' = \frac{\beta - \alpha}{\beta} L_1, \qquad C_1' = \frac{1}{L_1'\Omega_1^2},$$

$$L_2' = \frac{\alpha - \beta}{\alpha} L_1, \qquad C_2' = \frac{1}{L_2'\Omega_2^2}, \tag{13}$$

where

$$\Omega_1^2 = \frac{1}{2}\left[\omega_{11}^2 + \omega_{22}^2 + \{(\omega_{11}^2 - \omega_{22}^2)^2 + 4\omega_{12}^2\omega_{21}^2\}^{1/2}\right],$$

$$\Omega_2^2 = \frac{1}{2}\left[\omega_{11}^2 + \omega_{22}^2 - \{(\omega_{11}^2 - \omega_{22}^2)^2 + 4\omega_{12}^2\omega_{21}^2\}^{1/2}\right].$$

In the circuit of Fig. 2, the inductances and capacitances are positive according to Eqs. (13).

It is convenient to make the substitution

$$x_i = \gamma I_i', \qquad i = 1, 2.$$

The Kirchhoff equations for Fig. 2 are

$$\frac{d^2x_i}{dt^2} + \Omega_i^2 x_i + \mu f_i\left(x_1, x_2, \frac{dx_1}{dt}, \frac{dx_2}{dt}\right) = 0, \qquad i = 1, 2, \tag{14}$$

in which

$$L_1' f_1 = [(x_1 + x_2)^2 - 1]\left(\frac{dx_1}{dt} + \frac{dx_2}{dt}\right) = L_2' f_2 \,. \tag{15}$$

Equations (14) together with (15) are an extension of Van der Pol's equation.

For the case of μ very small, one may proceed directly to equations (7) after evaluating the integrals (6). If Ω_1 and Ω_2 are not commensurate (non-resonant case), the corresponding integrals (9) are

$$K_i(0) = \frac{1}{4\pi^2 L_i'} \int_0^{2\pi} \int_0^{2\pi} [(r_1 \cos\theta_1 + r_2 \cos\theta_2)^2 - 1](\omega_1 r_1 \sin\theta_1 + \omega_2 r_2 \sin\theta_2) \sin\theta_i \, d\theta_1 \, d\theta_2 \,, \qquad i = 1, 2 \tag{16}$$

so that equations (7) take the form

$$\frac{dr_1}{dt} = \frac{\mu}{8L_1'} r_1(4 - r_1^2 - 2r_2^2)$$

$$\frac{dr_2}{dt} = \frac{\mu}{8L_2'} r_2(4 - r_2^2 - 2r_1^2). \tag{17}$$

The singular points of (17) occur at (0,0), (2,0), (0,2), and $(2/3^{1/2}, 2/3^{1/2})$. They are respectively an unstable star point, a stable nodal point, a stable nodal point, and a saddle point. The corresponding limit cycles of (14) have the sets of amplitudes (0,0), (2,0), (0,2), and $(2/3^{1/2}, 2/3^{1/2})$ and are respectively unstable, stable, stable, and unstable.

The transient behavior of the amplitudes of the normal oscillations of (14) can be shown in the $r_1 r_2$ phase plane of the system (17). The differential equation of the phase plane trajectories is obtained by dividing the second Eq. (17) by the first. Thus

$$\frac{dr_2}{dr_1} = \frac{L_1'}{L_2'} \frac{r_2}{r_1} \frac{4 - r_2^2 - 2r_1^2}{4 - r_1^2 - 2r_2^2}.$$

For the case $L_1' = L_2'$, the solution is

$$(r_2^2 - r_1^2)^3 = 4\left[r_2^4 - r_1^4 + r_1^2 r_2^2\left(c - 2\ln\frac{r_2^2}{r_1}\right)\right], \tag{18}$$

c being the constant of integration. The trajectories are shown in Fig. 3.

If Ω_1 and Ω_2 are commensurate, that is, if there exist two integers, n_1 and n_2, such that

$$n_1\Omega_1 = n_2\Omega_2 \,,$$

equation (11) must be used to evaluate K_1 and K_2. The integrals are easily evaluated and yield four cases depending upon the ratio n_2/n_1. If the ratio is unity, the auxiliary set of differential Eqs. (7) contains the synchronous phase $\theta_{10} - \theta_{20}$. If the ratio is 1/2 or 3, the auxiliary set contains the synchronous phase $3\theta_{10} - \theta_{20}$ or $3\theta_{20} - \theta_{10}$. Any other ratio yields equations identical with Eqs. (17) describing the non-resonant case.

The location and stability of the limit cycles of the original system, as well as the transient behavior of the amplitudes of the normal modes, are strongly affected by the order of resonance. In the above example for the case of internal resonance of order

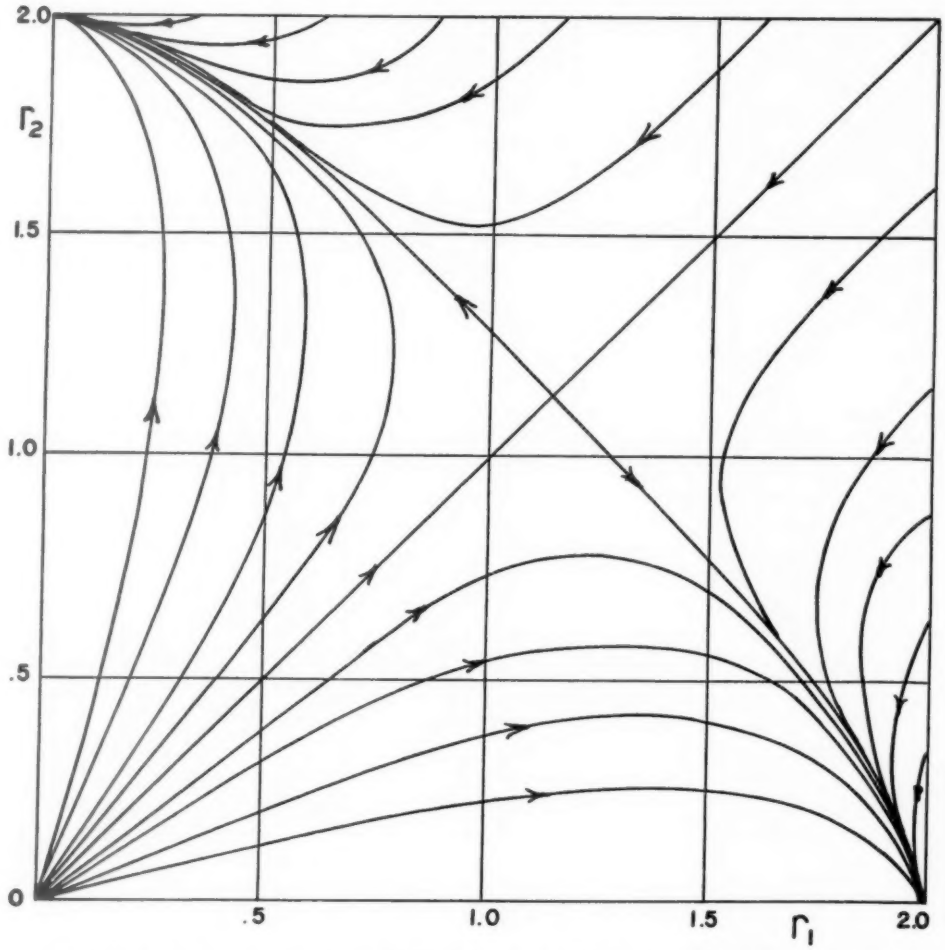

FIG. 3. Transient variation of the amplitudes of the normal modes of a doubly periodic system.

one-third ($n_1 = 3n_2$), there are four limit cycles. Three of the limit cycles have the respective amplitudes (0,0), (2,0), and (0,2), and correspond respectively to an unstable star point, a stable nodal point, and a stable nodal point in the auxiliary system. The remaining singular point is a saddle point; its location is a function of the synchronous phase

$$\theta_0 = 3\theta_{10} - \theta_{20} ,$$

and is shown in Figure 4.

Except for internal resonance of order one or one-third, the auxiliary equations are the same as for the non-resonant case, and hence the transient motion is not affected. The effect of resonance on the location and stability of limit cycles is complicated even

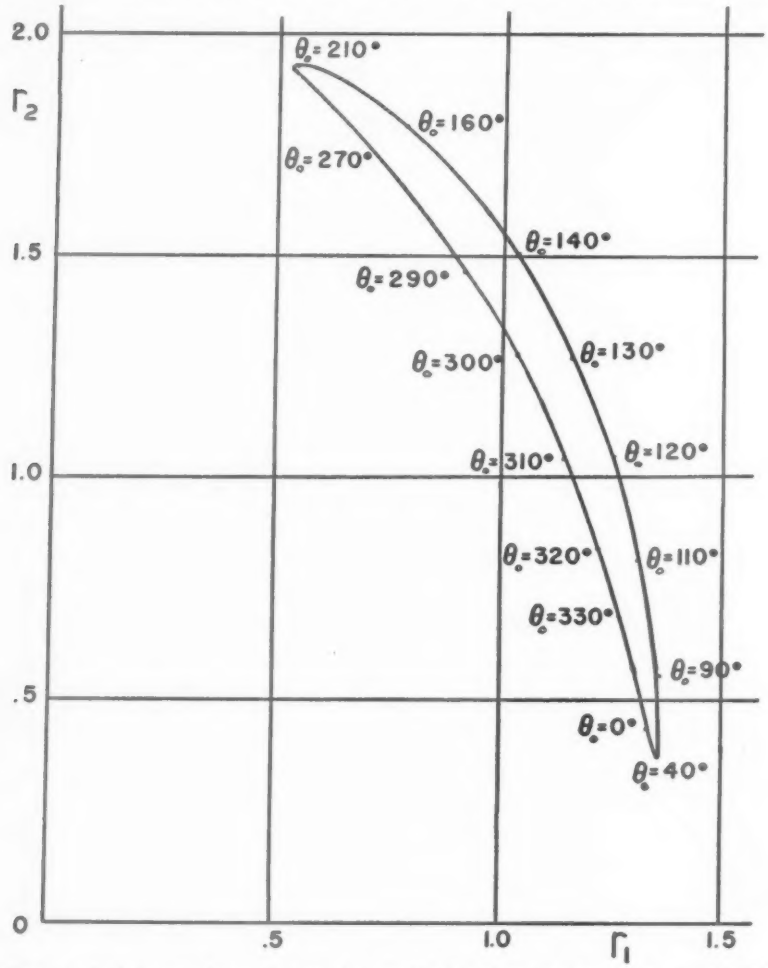

FIG. 4. Effect of synchronous phase, θ_0 , on the amplitudes of oscillation of the normal modes of a doubly periodic system.

in this simple case of a doubly periodic system, and will not be discussed in more detail here.

REFERENCES

1. H. Bateman, *Partial differential equations*, Dover Publications, New York, 1944, pp. 32-34.

2. N. Kryloff and N. Bogoliuboff, *Introduction to non-linear mechanics*, translation by S. Lefschetz, pp. 40-55, Princeton University Press, 1943.

PSEUDO CLOSED TRAJECTORIES IN THE FAMILY OF TRAJECTORIES DEFINED BY A SYSTEM OF DIFFERENTIAL EQUATIONS*

BY

L. A. MACCOLL

Bell Telephone Laboratories

1. Introduction. This paper is concerned with certain simple closed curves, here called pseudo closed trajectories, which play an important part in determining the topological properties of the family of trajectories (or characteristics) defined by a system of differential equations of the form

$$\frac{dx}{dt} = X(x, y), \qquad \frac{dy}{dt} = Y(x, y). \tag{1}$$

Some of these curves are considered in a rather incidental way in the writings of Poincaré.** However, the full concept of pseudo closed trajectories does not seem to have been discussed explicitly heretofore.

We assume that all of the variables in the equations (1) are real, that the functions X and Y are continuous in an open connected region R in the xy-plane, that these functions satisfy Lipschitz conditions locally throughout R, and that the real curves $X = 0$ and $Y = 0$, if they exist, have only simple intersections. We assume that these points of intersection have no point of condensation in the region R. We also assume that if the functions X and Y are not linear, they satisfy conditions which are sufficient to insure that the topological properties of the family of trajectories in a small neighborhood of any singular point (point of intersection of the curves $X = 0$, $Y = 0$) other than a center are the same as in the case obtained by replacing X and Y by their linear approximations. The last assumption is satisfied if X and Y are, for instance, of class C^2 in R.

In brief, our assumptions are simply the ones that are usually employed in discussions of the family of trajectories. The necessary information about the implications of the assumptions is readily available in the literature.†

Since the notion of a trajectory is a familiar one, we have used the term so far without any explanations. However, in order to avoid the danger of future ambiguities, it will be well now to define the term explicitly.

Let $x = \varphi(t)$, $y = \psi(t)$ be a solution of the system of equations (1). It is understood that the solution is defined over a certain maximum (finite or infinite) open interval $I^{(t)}$ on the t-axis.

As t describes $I^{(t)}$ once in the sense of increasing t, the point $(x, y) = (\varphi(t), \psi(t))$ describes a certain curve T in R. We call T a trajectory.

---

*Received Aug. 3, 1949.

**See, for instance, *Oeuvres de Henri Poincaré*, t. I, p. 56 et seq. Poincaré's considerations concerning these curves are limited to pseudo closed trajectories which have only exterior vertices or only interior vertices, and to the possibility of trajectories being asymptotic to such pseudo closed trajectories. Exterior and interior vertices are defined in the next section.

†A good discussion, covering most of the points of interest, is given in the recent book by Tricomi, *Equazioni differenziali*, Einaudi, Torino, 1948.

Any singular point is itself a trajectory, a point trajectory; and these are the only point trajectories. A trajectory T which is not a point trajectory is oriented, the positive sense on T being that in which the curve is described as t increases.

2. Pseudo closed trajectories. Under our assumptions we may have just four kinds of singular points: nodes, foci, centers and saddle points. Since the topological properties of the family of trajectories in the neighborhood of any one of these singular points are well known, it is unnecessary to review them extensively here.

We recall that a node or focus is said to be stable or unstable according as the oriented arcs of trajectories in the neighborhood proceed toward or away from the singular point.

In the neighborhood of a saddle point there are just two arcs of trajectories which proceed toward the saddle point. These arcs, considered for the moment as being unoriented, together with the saddle point itself, form a smooth arc. Also, there are just two arcs of trajectories which proceed away from the saddle point; and these, together with the saddle point, form a smooth arc. The two smooth arcs described have distinct tangents at the saddle point. These phenomena are illustrated in the typical Fig. 1.

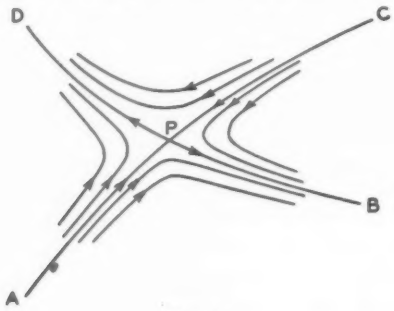

FIG. 1.

Elementary examples show that a trajectory may proceed *from* a saddle point *to* the same or a different saddle point.* We now assume that we have a case in which certain trajectories of this kind exist. Let the symbols T_1 , T_2 , $\cdots$ denote these trajectories.

Using a finite set of the trajectories T_i , and also using saddle points as necessary to connect these trajectories, we try to construct a closed curve Γ satisfying the following conditions:

C_1 : Γ is oriented, the orientation being that given by the orientations of the trajectories T_i contained in Γ.

C_2 : Γ is a simple closed curve.

Thus, referring to Fig. 1, C_1 implies that if AP is an arc of Γ, the succeeding arc is PB or PD. Also, if AP is an arc of Γ, C_2 implies that the arc CP does not belong to Γ.

We shall call a curve Γ satisfying the above conditions a *pseudo closed trajectory*, or briefly a *p.c.t.* There are families of trajectories which contain no p.c.t. On the other hand, familiar or easily constructed examples show that there are cases in which p.c.t. do exist. In all of the following we assume that we have a case of the latter kind.

*See, for instance, Andronow and Chaikin, *Theory of oscillations*, Princeton, 1949, p. 71.

We observe that different p.c.t. may have isolated points, or entire arcs, in common.

We shall call the saddle points which lie on a p.c.t. the *vertices* of the p.c.t. It is necessary to draw a distinction between two kind of vertices, which we call *exterior vertices* and *interior vertices*, respectively.

Let Γ be a p.c.t., and P one of its vertices. As a Jordan curve, Γ divides its point set complement in the plane into a exterior region and an interior region. In the neighborhood of P we have an arc of trajectory proceeding toward P, and an arc of trajectory proceeding from P, neither of these arcs belonging to Γ. If both of these arcs of trajectories are in the exterior region, we say that P is an exterior vertex; and if both of these arcs are in the interior region, we say that P is an interior vertex.* In the case of the p.c.t. Γ shown in Fig. 2, P_1 and P_2 are interior vertices, and P_3 is an exterior vertex.

3. The fundamental theorem concerning pseudo closed trajectories.

THEOREM. *Let Γ be a p.c.t. such that all points of the interior region determined by Γ belong to R. Let N_n, N_f, N_c, N_s, respectively, denote the numbers of nodes, foci, centers, and saddle points in the interior region, and let N_{iv} denote the number of interior vertices of Γ. Then we have the relation*

$$N_n + N_f + N_c - N_s = 1 + N_{iv}.$$

The proof of this theorem depends upon the following two familiar theorems.

T_1 : Let Γ be a simple closed plane curve consisting of a finite set of smooth arcs meeting at corners. As the running point describes Γ once in the counterclockwise sense, the tangent rotates through $+ 2\pi$ radians, it being understood that at an exterior (interior) corner the tangent rotates counterclockwise (clockwise) through the appropriate exterior (interior) angle.†

T_2 : The Poincaré index** of a simple closed curve which does not pass through any singular point, and which is such that the interior region determined by the curve belongs to R, is equal to the sum of the numbers of nodes, foci, and centers in the interior region, diminished by the number of saddle points in the interior region.

Let Γ be a p.c.t. with interior vertices $P_1, \cdots, P_k$. It is understood that these points are ranged in the order in which they are encountered when Γ is described in the sense of its natural orientation.

We choose points (not vertices) $Q_1', Q_1'', \cdots, Q_k', Q_k''$ on Γ so that the typical arc $Q_i'Q_i''$ of Γ contains P_i, and no other vertex. At each of the points Q_i', Q_i'' we draw the normal to Γ.

Now we construct a simple closed curve Γ^* as follows.

(1) For each i, we join a point on the exterior normal at Q_i' to a point on the exterior normal at Q_i'' by means of an arc of trajectory. The classical continuity theorems, together with the known behavior of the trajectories in the neighborhood of a saddle point, insure that this construction is possible, provided that the initial point of the arc of trajectory is taken sufficiently near to Γ.

---

*It is clear that it is impossible for one of these arcs to be in the exterior region and the other in the interior region.

†A rigorous modern proof of this intuitively evident theorem is given by J. Lifshitz, Boletin de la Sociedad Matematica Mexicana, 3, 21-25 (1946).

**Following the usual modern practice, we reckon the indices as the negatives of the indices actually defined by Poincaré.

(2) For $i = 1, \cdots, k-1$, we join a point on the interior normal at Q_i'' to a point on the interior normal at Q_{i+1}' by means of an arc of trajectory. If the initial point of the arc is taken sufficiently near Γ, this construction is possible, for the reasons mentioned above. Similarly, we join a point on the interior normal at Q_k'' to a point on the interior normal at Q_1' by an arc of trajectory.

(3) We complete Γ^* by joining the end points of the arcs of trajectories that we have drawn by short segments of the normals to Γ at the points Q_i', Q_i''.

For the sake of greater clarity, we emphasize the fact that the curve Γ^* is supposed to be constructed so that each arc of trajectory belonging to Γ^* lies very close to the corresponding arc of Γ (i.e. an arc $Q_i'Q_i''$, an arc $Q_i''Q_{i+1}'$, or the arc $Q_k''Q_1'$) throughout its length. This is possible, because we have constructed the curve so as to avoid the arcs of trajectories which proceed to and from the vertices of Γ, and which do not belong to Γ.

Fig. 2 illustrates the construction of Γ^* in a simple case.

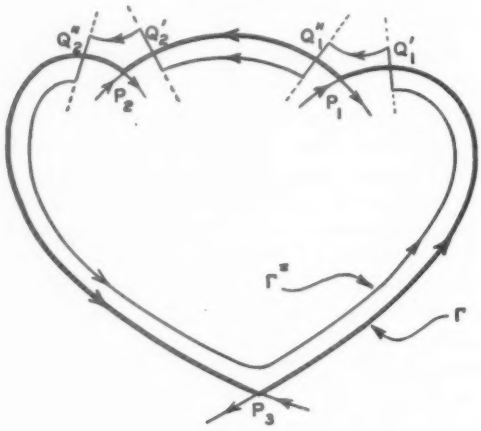

FIG. 2.

The curve Γ^* is closed, by its construction. A little consideration suffices to show that if the segments of normals to Γ belonging to Γ^* are taken sufficiently short, Γ^* is a simple closed curve, and that the interior region determined by Γ^* belongs to R. Also, under the same condition, Γ^* does not pass through any singular point; and the singular points contained in the interior region determined by Γ^* are the singular points contained in the interior region determined by Γ, plus the interior vertices of Γ. Therefore, by T_2 above, the index of Γ^* is $N_n + N_f + N_c - N_s - N_v$.

We now proceed to evaluate the index of Γ^* in a different way.

The index of Γ^* is the net number of revolutions performed by the field vector (X, Y) in the positive sense when the point (x, y) describes Γ^* once counterclockwise. The continuity of the functions X, Y implies that the contribution to the index, conceived in this way, coming from the short segments of normals to Γ is small in absolute value. Now consider the contribution coming from the arcs of trajectories belonging to Γ^*.

We note that at a point of Γ, other than a vertex, the field vector does not vanish,

and has the direction of the tangent to Γ. From this, and the continuity of X and Y, it follows at once that the contribution to the index of Γ^* coming from the arcs of trajectories belonging to Γ^* is nearly equal to the net number of revolutions performed by the tangent to Γ when the point of tangency describes Γ once counterclockwise. By T_1 above, this number of revolutions is $+1$.

The conclusions of the preceding two paragraphs imply that the index of Γ^* cannot be far different from $+1$. Hence, since the index is necessarily an integer, the index is $+1$ exactly.

When we equate the two expressions for the index that we have obtained, we have the conclusion of the theorem.

The preceding reasoning is based upon the hypothesis that Γ has at least two interior vertices. However, the same proof, with some slight and obvious modifications of the expository details, applies also when Γ has only one interior vertex.

If Γ has no interior vertices, we draw the normal to Γ at a point Q which is not a vertex. Then we construct a simple closed curve Γ^* out of a short segment of the interior normal and an arc of trajectory which joins the end points of the segment, and which lies very close to Γ throughout its length. Having constructed Γ^* in this way, we obtain the conclusion of the theorem by means of reasoning which is essentially similar to that used above.

4. Study of some simple cases. In the case of a p.c.t. that has no interior vertices, we have the relation

$$N_n + N_f + N_c - N_s = 1.$$

This is the same as the relation that we have in the case of a closed trajectory (in the ordinary strict sense). Consequently, the existence of a p.c.t. without interior vertices leads to essentially the same conclusions, in regards to the topological properties of the family of trajectories in the interior region, that the existence of a closed trajectory does. Since these conclusions have been thoroughly discussed in the literature, we shall not consider them here. Instead, we proceed to consider some simple cases in which we have p.c.t. with interior vertices.

Case 1.

Suppose that Γ is a p.c.t. with just one interior vertex P, and suppose that there are no centers, saddle points, or closed trajectories in the interior region determined by Γ. (It is understood that the interior region belongs to R.)

The fundamental theorem implies that there are just two singular points in the interior region. Either of these points, which we shall denote by the symbols S_1 , S_2 , may be a node or a focus. We shall see that when the types of these singular points are determined, there remain just two topological possibilities as to the structure of the family of trajectories in the interior region.

In the interior region there is one arc of trajectory which proceeds from P, and one arc of trajectory which proceeds to P. In the first subcase, illustrated by Fig. 3a, these arcs belong to the same trajectory T. The trajectory T and the saddle point P form a p.c.t. Γ' with just one vertex, which is exterior. The interior region determined by Γ' is contained in the interior region determined by Γ. The fundamental theorem implies that one of the points S_1 , S_2 , say the point S_1 , lies in the interior region determined

by Γ'. Then S_2 must lie in the complement of this region with respect to the interior region determined by Γ.

Suppose that the arcs of trajectories which proceed from and to P, respectively, do not belong to the same trajectory. Then the arc of trajectory proceeding from P cannot proceed to any singular point on Γ, for all of the arcs of trajectories proceeding to such singular points are otherwise accounted for. The arc of trajectory cannot approach asymptotically to a closed trajectory, for there are no such trajectories in the interior

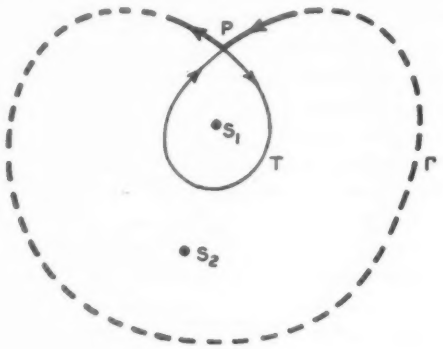

FIG. 3a.

region. Simple considerations, which we omit for the sake of brevity, show that the arc of trajectory cannot approach asymptotically to Γ. It follows, therefore, that the arc of trajectory must proceed to one of the points S_1 , S_2 , say the point S_1 . Similar reasoning shows that the arc of trajectory which proceeds to P must proceed from S_2 .

We have the situation illustrated in Fig. 3b, which is drawn on the basis of the

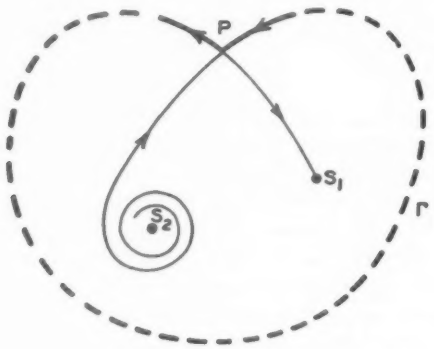

FIG. 3b.

assumption that S_1 is a node and S_2 a focus. The appropriate modifications of the figure for the cases in which S_1 and S_2 are of other types are obvious.

We note that in this second subcase S_1 and S_2 must be stable and unstable, respectively.

CASE 2.

Suppose that we have a p.c.t. Γ with just two interior vertices, P_1 and P_2. Suppose also that the interior region determined by Γ, which is understood to belong to R, contains no centers, saddle points, or closed trajectories. Then the fundamental theorem implies that there are just three singular points, S_1, S_2, and S_3, in the interior region. Any one of these may be either a node or a focus. For the sake of simplicity we draw

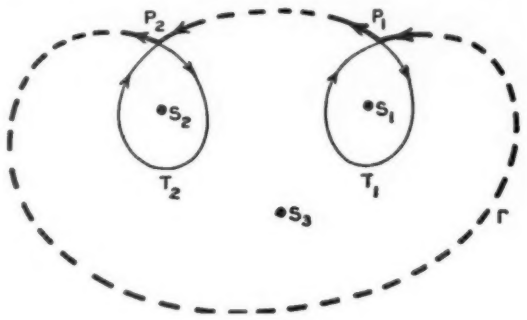

FIG. 4a.

the figures for the case in which all three of these singular points are nodes. The modifications of the figures for the other cases are obvious. We shall see that when the types of S_1, S_2, S_3 are determined there remain six different possibilities as to the topological properties of the family of trajectories in the interior region.

1. Consider the two arcs of trajectories, in the interior region, which proceed from P_1 and P_2, respectively. It may be that the arc proceeding from $P_1(P_2)$ and the arc of trajectory proceeding to $P_1(P_2)$ belong to the same trajectory $T_1(T_2)$. This is the situation illustrated in Fig. 4a. Reasoning which is similar to that employed in the preceding

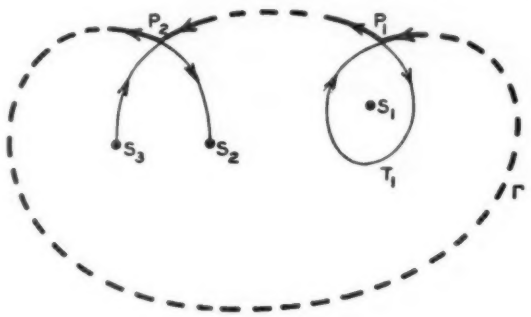

FIG. 4b.

case shows that if the subscripts are assigned suitably, the singular points S_1, S_2, S_3 must be located as shown in the figure.

2. It may be that the arc of trajectory proceeding from P_1 and the arc of trajectory proceeding to P_1 belong to the same trajectory T_1, but that the like is not true of the

arcs of trajectories proceeding from and to P_2 . Then one of the singular points S_i , say the point S_1 , must be located in the interior region determined by the p.c.t. Γ' formed by T_1 and P_1 . Also, we see that the arc of trajectory proceeding from P_2 must proceed to S_2 or S_3 say to S_2 , and that the arc of trajectory proceeding to P_2 must proceed from S_3 . This situation is illustrated in Fig. 4b.

3. It may be that the arc of trajectory which proceeds from P_1 proceeds to P_2 , and that the arc of trajectory which proceeds from P_2 proceeds to P_1 . This situation is illustrated in Fig. 4c. The fundamental theorem implies that the points S_1 , S_2 , S_2 ,

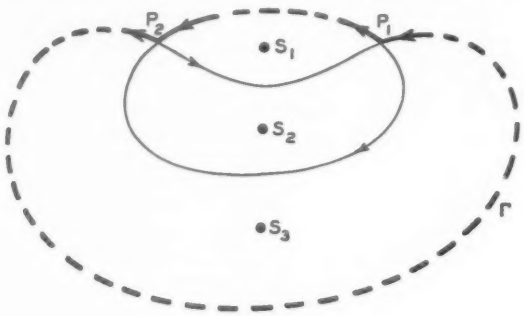

FIG. 4c.

with a suitable assignment of the subscripts, must be located with respect to the trajectories as shown.

4. It may be that the arc of trajectory which proceeds from P_1 proceeds to one of the singular points S_i , say the point S_1 , and that the arc of trajectory which proceeds from P_2 proceeds to P_1 . Then it is easily seen that the arc of trajectory which proceeds to P_2 must proceed from another singular point, say S_2 . The third singular point must be located as shown in the illustrative Fig. 4d.

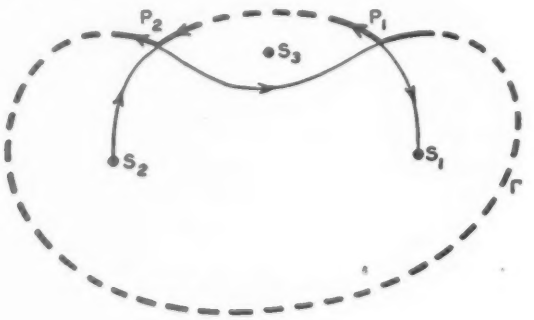

FIG. 4d.

5. In each of the four preceding subcases we have at least one arc of traectory which proceeds from one of the vertices P_1 , P_2 to the same or the other vertex. Now suppose that this condition is not satisfied. Then the arcs of trajectories which proceed from P_1 and P_2 must proceed to the same or different ones of the points S_i . Suppose that the

two arcs proceed to the same singular point S_3 . It is clear that the arcs of trajectories proceeding to P_1 and P_2 , respectively, must lie on opposite sides of the arc $P_1S_3P_2$. Consequently, these arcs of trajectories must proceed from singular points S_1 and S_2 , which are located as shown in the illustrative Fig. 4e.

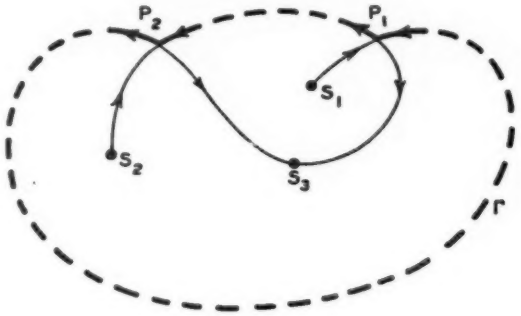

FIG. 4e.

6. Suppose, on the contrary, that the arcs of trajectories proceeding from P_1 and P_2 proceed to S_1 and S_2 , respectively. Then it is easily seen that the arcs of trajectories proceeding to P_1 and P_2 must proceed from S_3 . This subcase gives us a variant of Fig. 4e which we do not need to exhibit.

It is readily seen, in the light of the familiar Poincaré theory of families of trajectories, that our discussion of the cases we have considered is essentially complete topologically.

SOME RECENT INVESTIGATIONS IN THE THEORY OF HYDRODYNAMIC STABILITY*

BY

J. R. FOOTE (*Iowa State College*)

AND

C. C. LIN (*Massachusetts Institute of Technology*)

1. **Introduction.** In most of the existing investigations of the stability of two-dimensional parallel flows [1], two essential restrictions have been imposed: (1) at least one solid boundary is present, and (2) the velocity distribution is a monotone curve. It has been possible to treat the case of symmetrical velocity distributions through a channel by imposing the boundary condition at the axis of symmetry and therefore fulfilling the above conditions.

There is a large class of parallel flows where no solid boundary is present. These are the problems of "free boundary layers", such as the flow field in the jet, the wake, and the mixing region between two streams. It may be expected at once from physical considerations that the influence of the "viscous solutions" of the disturbance equation should not occur, because their influence is known to be chiefly of the nature of a boundary layer phenomenon in the immediate neighborhood of the solid boundary. In the case of the mixing region between two streams, this fact can also be easily seen mathematically. The exponential nature of the "viscous solutions" is well known, and this behavior is definitely incompatible with the condition that the disturbances must have finite amplitudes at large distances. The effect of viscosity must therefore be found from the second order approximation of the "inviscid solutions." The viscous influence is therefore restricted to the immediate neighborhood of the critical layer, and the general stability characteristics of such flows should therefore be very much different from those previously investigated.

With these ideas in mind, Lin suggested the investigation of the stability characteristics of such flows to M. Lessen [2] (using numerical methods) and to P. Chiarulli [3] (using analytical methods). Lessen made detailed calculations for one profile with the help of Dr. L. H. Thomas of I. B. M. of New York City. The results show that the effect of viscosity is always to stabilize the flow even for very large Reynolds numbers, contrary to Heisenberg's criterion for cases with solid boundary. The critical Reynolds number is found to be very low, as may be expected from the presence of the point of inflection in the velocity distribution and the limited extent of the effect of viscosity. Chiarulli obtained results for general velocity distributions of the mixing type. It appears from his analysis that the stabilizing effect of viscosity is probably a general characteristic of such flows.

For jets and wakes it is not obvious that the above discussions apply. In the case of symmetrical jets and wakes, one may apply the usual type of analysis and reach the above conclusions concerning the origin of the viscous effects (see Sec. 2). The case of unsymmetrical distributions is in general more difficult, because there are usually

*Received Aug. 21, 1949. A part of this paper is based on a thesis submitted by J. R. Foote in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology.

two critical points in the flow field. The purpose of the present paper is to clarify the nature of the asymptotic solutions in cases with two critical points. It will be shown by a slight extension of some work of Wasow [4] that to preserve the validity of one asymptotic expression for the solutions, a path must be taken *below* the critical point in the region where the slope of the velocity curve is positive, and *above* the critical point in the region with negative slope. Application of this result to the cases of unsymmetrical jets and wakes again shows that the "viscous solutions" must be rejected and that the effect of viscosity must come from the second order approximation of the "inviscid solutions".

The remaining part of the paper is devoted to an investigation of the inviscid approximation, with special reference to the stability of jets, wakes, and the zonal winds in the atmosphere. It is shown that, in all cases, a necessary and sufficient condition for the existence of (non-trivial) neutral and self-excited disturbances is the occurrence of the maximum or minimum of vorticity. This is true even in the case of zonal winds where the point of inflection of the profile does not have the usual significance, and the usual proof of sufficiency does not apply.

It is further shown that for unsymmetrical velocity distributions the neutral disturbance may give rise to a Reynolds shear which is opposite in sign to the velocity gradient of the mean flow. In other words, these ordered oscillations, unlike fully developed turbulent motions, can transmit momentum from a region of lower mean velocity to a region of higher mean velocity

2. Asymptotic solutions and boundary value problems. The derivation of the linearized stability equation

$$(w - c)(\varphi'' - \alpha^2\varphi) - w''\varphi = -\frac{i}{\alpha R}(\varphi^{IV} - 2\alpha^2\varphi'' + \alpha^4\varphi), \tag{1}$$

and the explanation of symbols can be found in [1]. We consider jet profiles extending to infinity in both directions. Boundary conditions for these profiles may be written

$$\lim_{y\to\pm\infty}\left[\frac{\varphi'(y)}{\varphi(y)}\right] = \mp\alpha,$$

$$\lim_{y\to\pm\infty} u = 0, \qquad \lim_{y\to\pm\infty} v = 0. \tag{2}$$

For the symmetric profile one may use one of these conditions together with one of the following:

$$\varphi(0) = 0, \qquad \text{if } \varphi(y) \text{ is odd}, \tag{2a}$$

$$\varphi'(0) = 0, \qquad \text{if } \varphi(y) \text{ is even}. \tag{2b}$$

It has been shown by Wasow [4] that the asymptotic solutions of (1) may be obtained by formal expansions. In Sec. 6 we shall extend his result slightly and discuss how these solutions can be applied to profiles having more than one critical point. It will be shown that there are asymptotic solutions valid in a region defined as follows. The critical points where $w(y) = c$ shall be denoted by y_c and y_0, with $Re[w'(y_c)] > 0$ and $Re[w'(y_0)] < 0$. Through the point y_c there are three lines C_k, $k = 1, 2, 3$, defined by

$$ReQ_c(y) = Re\left\{\int_{y_c}^{y} [i(w - c)]^{1/2}\, dy\right\} = 0, \tag{3}$$

with three similar lines C_k^*, $k = 1, 2, 3$, through y_0 (Fig. 1). Consider a path Γ extending

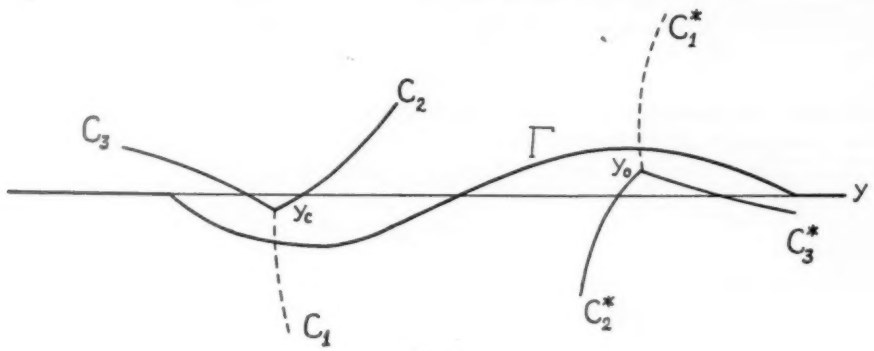

FIG. 1.

from $y = -\infty$ to $y = +\infty$, and crossing only one each of the lines C_k and C_k^*. A neighborhood of such a path Γ is a region of validity of the asymptotic solutions.

The explicit expansions of the solutions are:

$$\begin{aligned}
\varphi_1 &= \varphi_1^{(0)} + \varphi_1^{(1)}(\alpha R)^{-1} + \varphi_1^{(2)}(\alpha R)^{-2} + \cdots \\
\varphi_2 &= \varphi_2^{(0)} + \varphi_2^{(1)}(\alpha R)^{-1} + \varphi_2^{(2)}(\alpha R)^{-2} + \cdots \\
\varphi_3 &= \exp\,[-(\alpha R)^{1/2} Q_c(y)]\{(w-c)^{-5/4} + O(\alpha R)^{-1/2}\} \\
\varphi_4 &= \exp\,[+(\alpha R)^{1/2} Q_c(y)]\{(w-c)^{-5/4} + O(\alpha R)^{-1/2}\}
\end{aligned} \tag{4}$$

The notation agrees with that used in [1]. The solutions φ_3 and φ_4 respectively become infinite exponentially as $y \to -\infty$ and $+\infty$. It is then clear that $c_3 = c_4 = 0$ in the general solution

$$\varphi = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4$$

if we apply the boundary conditions at $y = \pm\infty$, for either the symmetric or the unsymmetric profile. Thus the effect of viscosity would appear to enter only through the higher approximations for φ_1 , φ_2 . This should be discussed further for the symmetric profile since (2a) or (2b) may be used with one of the conditions (2). Then we are using only half the profile, so that φ_4 does not diverge and may be included:

$$\varphi = c_1\varphi_1 + c_2\varphi_2 + c_4\varphi_4 .$$

For case (2a) the determinantal eigenvalue condition is then

$$\begin{vmatrix} \varphi_{11}' - \alpha\varphi_{11} & \varphi_{21}' - \alpha\varphi_{21} & 0 \\ \varphi_{12} & \varphi_{22} & \varphi_{42} \\ \varphi_{12}'' & \varphi_{22}'' & \varphi_{42}'' \end{vmatrix} = 0, \tag{5}$$

where we have put $y_1 = -\infty$, $y_2 = 0$, and $\varphi_{11}' = \varphi_1'(y_1)$, $\varphi_{12} = \varphi_1(y_2)$, and so on. How-

ever, we shall show that the viscous effects are still not introduced by the "viscous solutions."

Suppose $\Phi = k_1\varphi_1 + k_2\varphi_2$ is a solution satisfying the inviscid equation

$$(w - c)(\varphi'' - \alpha^2\varphi) - w''\varphi = 0 \tag{6}$$

to the first approximation and satisfying the boundary condition at y_1. Taking this linear combination of the first two columns in (5) and replacing the first column by the result and simplifying, we obtain the eigenvalue condition as

$$\Phi_2 = \Phi_2'' \frac{\varphi_{42}}{\varphi_{42}''}. \tag{7}$$

Using (4) we find expressions for Φ_2 and $\varphi_{42}/\varphi_{42}''$. Then (7) becomes

$$\begin{aligned} &\Phi_2^{(0)} + \frac{\Phi_2^{(1)}}{\alpha R} + \cdots \\ &\qquad = \left[\Phi_2^{(0)''} + \frac{\Phi_2^{(1)''}}{\alpha R} + \cdots\right] \frac{1}{i(w - c)_2 \alpha R}\,[1 + \cdots]. \end{aligned} \tag{8}$$

Neglecting terms of order $0(\alpha R)^{-1}$, (8) yields:

$$\Phi_2^{(0)} = 0.$$

Then it follows from (6) that $\Phi_2^{(0)''}$ also vanishes. If we now include terms of order $O(\alpha R)^{-1}$, (8) yields:

$$\Phi_2^{(0)} + \frac{\Phi_2^{(1)}}{\alpha R} = 0,$$

and again there is no contribution to viscosity from φ_4. This process can be continued with the aid of the recursion formulae for φ_1 and φ_2[1].

We can perform the same steps with case (2b) and find the condition corresponding to (7) as

$$\Phi_2' = \Phi_2''' \frac{\varphi_{42}'}{\varphi_{42}'''}$$

Here $\varphi_{42}'/\varphi_{42}'''$ has the same form as in (8), but (6) gives

$$\Phi_2^{(0)'''} = \left(\frac{w''}{w - c}\right)_2' \Phi_2^{(0)} + \left(\alpha^2 + \frac{w''}{w - c}\right)_2 \Phi_2^{(0)'}.$$

Now w'' and $w - c$ are even functions of y, so their ratio is even and the derivative shown is odd and hence zero at $y_2 = 0$. Hence $\Phi_2^{(0)'''}$ vanishes with $\Phi_2^{(0)'}$, and the calculations proceed in exactly the same way. Therefore we conclude that *the effect of viscosity enters the stability problem for the jet profile (symmetric or unsymmetric) only through the higher approximations to the "inviscid solutions" of the stability equation* (1). This is an extension of the conclusion reached in the infinite jet (mixing problem) and is in marked contrast to other stability problems previously investigated.

We note the following important facts: (a) the only usable solutions of (1) tend, as $\alpha R \to \infty$, to a fundamental system of (6), which may be obtained from (1) by formally passing to the limit of infinitely large Reynolds number; (b) The stability problem for

(1) (expressed in the eigenvalue condition) tends to the stability problem of (6) as $\alpha R \to \infty$. These facts give importance to a study of the inviscid equation.

3. The inviscid equation. The inviscid equation (8) is essentially the vorticity equation for a perfect fluid. The term $-w''(y)$ is known to represent the gradient of vorticity [1]. In other problems of stability of parallel flows (see Sec. 5), the gradient of vorticity also enters the problem, but it is not necessarily related to the main velocity distribution $w(y)$ in this simple manner. We shall therefore consider the equation

$$[w(y) - c](\varphi'' - \alpha^2\varphi) + g(y)\varphi = 0, \tag{9}$$

where $g(y)$ is the gradient of vorticity, and is to be replaced by $-w''(y)$ in the simple case. It is obvious that the results of Sec. 2 are applicable without modification, because they depend only on the coefficient $w - c$ of φ''.

Solution of (9) by means of power series gives the fundamental system*

$$\begin{aligned} \varphi_1(y) &= (y - y_c) + \cdots \\ \varphi_2(y) &= 1 + \cdots - \frac{g(y_c)}{w'(y_c)}\varphi_1 \log (y - y_c) \end{aligned} \tag{10}$$

in the neighborhood of $y = y_c$, and a similar system in the neighborhood of the critical point $y = y_0$. The only difference lies in the path for the determination of the logarithmic function. The path is to be taken below the point y_c, but above the point y_0 (see Fig. 1). Thus, in the case of self-excited disturbances, the inviscid equation holds all along the real axis. But in the case of damped disturbances it fails over portions of the real axis in the neighborhood of y_c and y_0. The general solution of (9) in the neighborhood of y_c is

$$\varphi = A\varphi_1 + B\varphi_2 = A\varphi_1 + B\varphi_{10} - B\frac{g_c}{w_c'}\varphi_1 \log (y - y_c), \tag{11}$$

where φ_{10} is the power series part of φ_2 and A and B are complex constants.

It is found most convenient to study the inviscid equation by examining the Reynolds shear stress, as was first done by Tollmien [5]. We shall now generalize his analysis from several points of view and then apply the results to various types of velocity distributions.

Assuming the existence of the solutions, we shall establish three general relations. The actual existence problem will be discussed later.

(i) *The Reynolds shear stress is*

$$\tau = -\rho\overline{u'v'} = \rho\frac{\alpha^2}{2}e^{c_i t}W(\varphi_r, \varphi_i), \tag{12}$$

where ρ is the density of the fluid, u' and v' are the real disturbance velocities, the averaging process is taken with respect to x, and $W(\varphi_r, \varphi_i)$ is the Wronskian $\varphi_r\varphi_i' - \varphi_i\varphi_r'$ of the real and imaginary parts $\varphi_r(y)$ and $\varphi_i(y)$ of $\varphi(y)$ for real values of y. Tollmien discussed this relation only for neutral solutions with $c_i = 0$.

(ii) *The rate of change of the Reynolds shear stress τ along the real axis is given by*

$$\frac{d\tau}{dy} = -\rho\frac{\alpha^2}{2}e^{2c_i t}c_i g\,|\,\varphi\,|^2/|\,w - c\,|^2. \tag{13}$$

*This is a slight generalization of Tollmien's result [5].

It is obvious that both the above results are applicable only for parts of the real axis where φ is regular.

(iii) *In the case of the neutral disturbance, there is a discontinuity of the Reynolds shear at the critical point of amount*

$$[\tau] = \tau(y_c + 0) - \tau(y_c - 0) = -\frac{\alpha^2}{2}\pi\rho \mid \varphi_c \mid^2 g_c/\mid w_c' \mid. \tag{14}$$

All these results can be obtained by straightforward calculations. The real disturbance velocities u' and v' (with real values of y) are given by

$$u' = Re(\psi_y'), \qquad v' = Re(-\psi_x'),$$

with the disturbance stream function

$$\psi'(x, y, t) = \varphi(y)\, e^{i\alpha(x-ct)}.$$

Thus, the average

$$\overline{u'v'} = \lim_{x\to\infty} \frac{1}{x}\int_0^x u'v'\, dx$$

is easily seen to be

$$\overline{u'v'} = \frac{\alpha^2}{2} e^{2c_i t}(\varphi_i\varphi_r' - \varphi_r\varphi_i'),$$

from which (12) follows immediately. One may also write

$$\varphi_i\varphi_r' - \varphi_r\varphi_i' = \frac{1}{2i}(\varphi\bar{\varphi}' - \bar{\varphi}\varphi'),$$

so that

$$W(\varphi_r, \varphi_i) = \frac{i}{2}(\varphi\bar{\varphi}' - \bar{\varphi}\varphi'). \tag{15}$$

It should be noted that the conjugate signs are formal conjugates for real values of y, i.e.,

$$\bar{\varphi} = \varphi_r(y) - i\varphi_i(y), \qquad \bar{\varphi}' = \varphi_r'(y) - i\varphi_i'(y).$$

From (15) one can easily derive (13) as follows. Rewriting (9) in the form

$$\varphi'' - \alpha^2\varphi + \frac{g(y)}{w(y) - c}\varphi = 0, \tag{9a}$$

multiplying it by $\bar{\varphi}$ and subtracting from it the complex conjugate equation, one obtains

$$\frac{d}{dy}(\varphi'\bar{\varphi} - \bar{\varphi}'\varphi) = -\frac{\mid \varphi \mid^2}{\mid w - c \mid^2} 2c_i g_i,$$

or

$$\frac{dW}{dy} = -\frac{\mid \varphi \mid^2}{\mid w - c \mid^2} c_i g. \tag{16}$$

By combining this relation with (12), the relation (13) is obtained at once.

To obtain (14) it is only necessary to calculate the jump of W, which in turn is determined from the jump of φ', since φ is continuous across the singular point. Thus, from (15),

$$[W] = \frac{i}{2}\{\varphi[\overline{\varphi'}] - \bar{\varphi}[\varphi']\}, \qquad \text{for } y = y_c .$$

From (10) and (11) it is clear that

$$\varphi_c = B, \qquad [\varphi'] = -B\frac{g_c}{|w_c'|}i\pi,$$

since the angle of $y - y_c$ increases by π when y passes from $y_c - 0$ to $y_c + 0$ if $w_c' > 0$ and decreases if $w_c' < 0$*. Thus,

$$[W] = -\pi\frac{g_c}{|w_c'|}|\varphi_c|^2. \tag{17}$$

By combining (17) and (12), (14) follows at once.

We shall now prove the following theorem giving a necessary condition for the existence of self-excited and neutral disturbances. Theorems on sufficiency of conditions will be discussed in the next sections.

THEOREM I. *If the disturbance must satisfy homogeneous boundary conditions at each of two points y_1 and y_2, the condition $g(y_s) = 0$ for some y_s in the interval (y_1, y_2) is necessary for the existence of self-excited disturbances, and of neutral disturbances with $w = c$ at some interior point of the interval (y_1, y_2).*

In the case $g(y) = -w''(y)$, neutral disturbances with $w \neq c$ in the closed interval (y_1, y_2) are also excluded by the condition $g(y) \not\equiv 0$ except possibly the trivial solution with $\alpha = 0$, $\varphi = (w - c)$, (which must be examined with the boundary conditions).

PROOF: The first part can be easily proved by examining the variation of τ for real values of y. At the points y_1 and y_2, τ vanishes. If a self-excited disturbance exists, and $g(y)$ does not vanish between y_1 and y_2, then (13) shows that τ varies monotonically in between, and hence cannot possibly fulfill the boundary conditions.

For a neutral disturbance, τ is a constant except for the jumps given by (14). To satisfy the boundary conditions, one must have

$$\sum [\tau] = 0, \tag{18}$$

summation being taken for all the jumps. Since $w_{\min} < c < w_{\max}$, there is at least one jump. If there is only one, we see at once that $g_c = 0$, or alternatively $\varphi_c = 0$. We must show that the second alternative is impossible. Now from (10) and (11), it is clear that $\varphi_c = 0$ implies that

$$\varphi = (w - c)f, \tag{19}$$

where f is a regular function. If we now multiply (9a) by $\bar{\varphi}$ and integrate between y_1 and y_c, we obtain

$$\int_{y_1}^{y_0}\{|\varphi'|^2 + \alpha^2|\varphi|^2\}\,dy = \int_{y_1}^{y_c} g(w - c)|f|^2\,dy \geq 0. \tag{20a}$$

*The result therefore also applies to the neighborhood of y_0.

Similarly,

$$\int_{y_c}^{y_2} g(w - c) \mid f \mid^2 dy \geq 0. \tag{20b}$$

Since $w - c$ definitely has opposite signs in the two intervals (y_1 , y_c) and (y_c , y_2), it is obvious that we have a contradiction if the sign of g remains unchanged.

If there is more than one point where $w = c$, either (i) all the φ_c's are zero, or (ii) some of the g_c's vanish or have opposite signs. The first case can be excluded by an argument similar to the above. The other case insures the vanishing of $g(y)$ within (y_1 , y_2).

To establish the second part of the theorem, let us note that we again have (19) when $w \neq c$ in (y_1 , y_2). We may rewrite (9) as

$$\frac{d}{dy}\{(w - c)^2 f'\} - \alpha^2(w - c)^2 f + (w'' + g)(w - c)f = 0. \tag{9b}$$

Multiplying this equation by $\bar{f}$ and integrating between y_1 and y_2 , we obtain

$$\int_{y_1}^{y_2} (w - c)^2\{\mid f' \mid^2 + \alpha^2 \mid f \mid^2\} dy = \int_{y_1}^{y_2} (w'' + g)(w - c) \mid f \mid^2 dy. \tag{21}$$

This relation does not allow much to be inferred, except in the important case $g = -w''$. In this case, we must have

$$\alpha = 0, \qquad f = \text{const.}, \tag{22}$$

which proves our statement. This method of proof is well-known [6], but we include the details here to show why it does not work for general $g(y)$.

4. Wakes and jets. We shall now investigate more closely several typical problems. In this section, we shall study the stability of two-dimensional wakes and jets, symmetrical and unsymmetrical. For these problems, $g(y) = -w''(y)$. In the next section, we shall study the stability of zonal winds over a surface of revolution.

(a) *Symmetrical jets.* In this case, the neutral disturbance can occur only when the wave velocity c is equal to the flow velocity at the point of inflection, except for the trivial case given in Theorem I. The Reynolds shear is zero everywhere, and there is no transfer of energy between the main flow and the disturbance.

For self-excited oscillations, (13) leads to a shear distribution shown in Fig. 2. One notes that $\tau > 0$ wherever $w'(y) > 0$. The energy is transferred from the main flow to the disturbance by the Reynolds shear, as it does in fully developed turbulent motion. For damped disturbances, the inviscid theory is applicable only to certain parts of the flow field. But (13) shows clearly that there will be regions where the shear τ has an opposite sign to the velocity gradient $w'(y)$; the damped disturbance feeds energy into the main flow in these regions.

We shall see that for unsymmetrical jets, τ and $w'(y)$ may have opposite signs even for the neutral disturbance.

We have yet to prove the existence of neutral and self-excited disturbances. This may be done in the usual manner for the even solution satisfying the boundary conditions $\varphi'(0) = 0$ and $\varphi \to 0$ as $y \to 0$. The proof given by Friedrichs [6] can be adopted without change, since $w(y)$ satisfies the boundary conditions required of $\varphi(y)$. The existence of self-excited disturbances can be proved by following the method first used

by Tollmien [5] and slightly extended by Lin [1, Part 2]. The convergence of the integrals involved can be easily verified.

The trivial solution $\varphi = w - c$ for $\alpha = 0$ is valid for *all* values of c, if one merely requires that the disturbance vanishes at infinity, for (u', v') are proportional to $(\alpha\varphi, \varphi')$.

(b) *Unsymmetric jets.* The existence of neutral and self-excited disturbances for such cases can best be seen from a continuity argument. Thus, if one decomposes the velocity distribution $w(y)$ into an even part $w_1(y)$ and an odd part $w_2(y)$, one may consider it as a member of the family $w_1(y) + \epsilon w_2(y)$. When ϵ is zero, self-excited disturbances exist for $0 < \alpha < \alpha_s$, say. When ϵ changes, the eigenvalue must change in a continuous

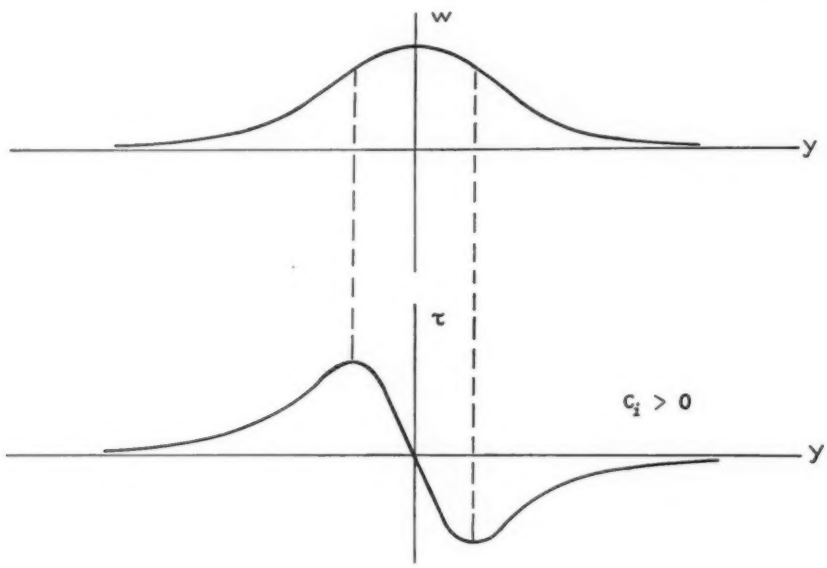

FIG. 2.

manner. Thus, there would still be a finite range of α for which the motion is self-excited and the end points of which correspond to neutral disturbances. But the range might shrink to zero when ϵ is fairly large. Thus, at least for velocity distributions not far from the symmetrical case, we have self-excited and neutral disturbances. One notes, however, that $\varphi = (w - c)$ is always a solution for $\alpha = 0$.

The shear stress distribution is very interesting in the present case. For definiteness, let the points where $w'' = 0$ be denoted by y_s and y_t with $y_s < y_t$ and $w_s > w_t$. For the neutral disturbance, it is then necessary that

$$w_s > c > w_t ,$$

because this is the only way to make $w_s''w_t'' < 0$. (See (14) and (18)). The distribution of stress τ is as shown in Fig. 3 (dotted). The stress suddenly jumps from zero to a finite value at y_s, stays constant, and then jumps back to zero at y_0. For the self-excited disturbance, (13) gives the distribution shown by the full line of Fig. 3. The maximum

stress occurs at y_s and the minimum at y_t . One notes that in either case there is a part where τ and $w'(y)$ are of opposite signs. The net transfer of energy in the neutral case is, however, zero, because

$$\int_{y_c}^{y_0} \tau w'(y)\, dy = \tau \int_{y_c}^{y_0} w'(y)\, dy = 0.$$

If $w_s < w_t$, the sign of τ is essentially negative. In either case, the shear stress in the case of the neutral disturbance transfers momentum from the part with lower velocity at the point of inflection to the part with higher velocity.

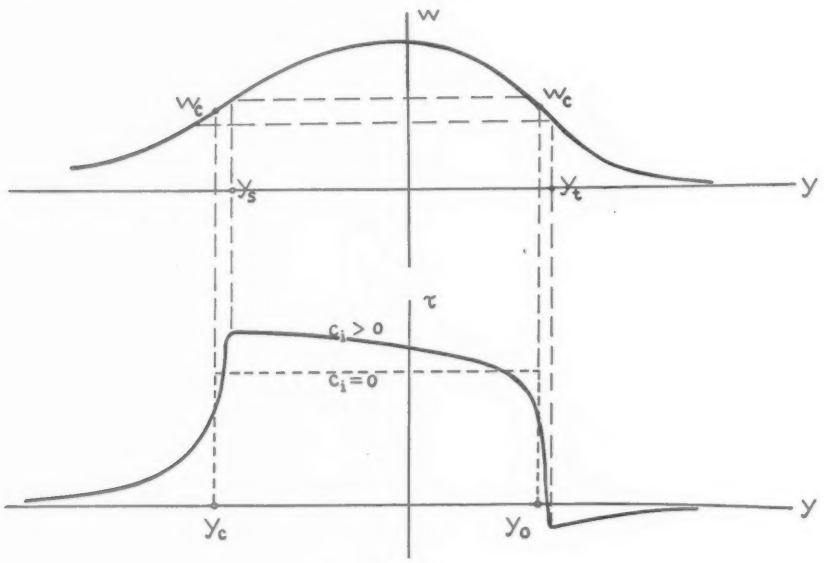

FIG. 3.

5. Instability of zonal winds*. We shall now apply our results to the motion of a thin layer of fluid over the surface of a spinning body of revolution, say, the earth. These results may be applied to stability of the zonal winds in the atmosphere. Discussions of such problems have been made by H. L. Kuo [8], largely by considering plane motions. With the help of present results, we can make a more thorough investigation of the problem.

Let us take an orthogonal coordinate system (φ, σ, z), where φ is the longitude, σ is the distance from the equator along a meridian positive towards the north pole, and z is the distance perpendicular to the surface. Consider two-dimensional motion of a thin layer of fluid in the surfaces of constant z. We may put $z = 0$ throughout the investigation. Let u and v be the linear velocity components relative to the earth along

*This section was added by Lin, after the rest of the paper had been completed. For related discussions of vortex motions in the atmosphere, see a recent paper by Rossby [7].

the directions of increasing φ and σ, and let ζ be the component of absolute vorticity in the direction of the z-axis. Then the vorticity equation

$$\frac{\partial \zeta}{\partial t} + \frac{u}{r}\frac{\partial \zeta}{\partial \varphi} + v\frac{\partial \zeta}{\partial \sigma} = 0, \tag{23}$$

where r is the distance from the axis of rotation, and

$$\zeta = \zeta_n + \frac{1}{r}\left\{\frac{\partial v}{\partial \varphi} - \frac{\partial}{\partial \sigma}(ru)\right\}, \tag{24}$$

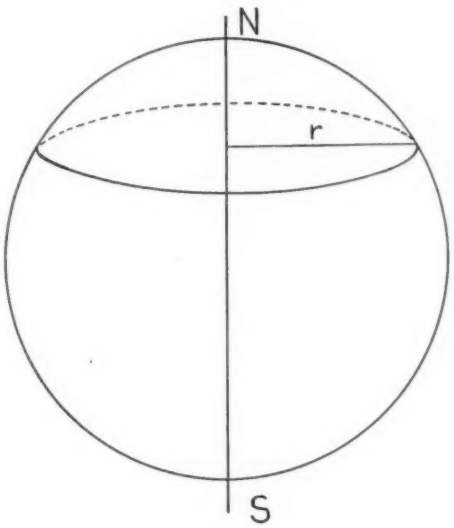

FIG. 4.

with ζ_n as the component of vorticity of spinning normal to the surface of the earth. The equation of continuity is

$$\frac{\partial u}{\partial \varphi} + \frac{\partial}{\partial \sigma}(rv) = 0. \tag{25}$$

Let us consider a motion slightly disturbed from a pure zonal motion

$$u = u_0(\sigma), \qquad v = 0. \tag{26}$$

The perturbations u', v' and ζ' then satisfy the linearized equations

$$\begin{aligned} &\frac{\partial u'}{\partial \varphi} + \frac{\partial}{\partial \sigma}(rv') = 0 \\ &\frac{\partial \zeta'}{\partial t} + \frac{u_0}{r}\frac{\partial \zeta'}{\partial \varphi} + v'\frac{d\zeta_0}{d\sigma} = 0, \end{aligned} \tag{27}$$

with

$$\zeta_0 = \zeta_n - \frac{1}{r}\frac{d}{d\sigma}(ru_0)$$

$$\zeta' = \frac{1}{r}\left\{\frac{\partial v'}{\partial \varphi} - \frac{\partial}{\partial \sigma}(ru')\right\}. \tag{28}$$

It is now possible to set up an "equivalent" plane motion by defining the following quantities:

$$x_* = \varphi, \qquad y_* = \int_0^\sigma \frac{d\sigma}{r}, \qquad w_* = \frac{u_0}{r}$$

$$u'_* = ru', \qquad v'_* = rv', \qquad \zeta'_* = r^2\zeta'. \tag{29}$$

Then

$$\frac{\partial u'_*}{\partial x_*} + \frac{\partial v'_*}{\partial y_*} = 0$$

$$\frac{\partial \zeta'_*}{\partial t} + w_* \frac{\partial \zeta'_*}{\partial x_*} + v'_* \frac{d\zeta_0}{dy_*} = 0 \tag{30}$$

$$\zeta'_* = \frac{\partial v'_*}{\partial x_*} - \frac{\partial u'_*}{\partial y_*}$$

Thus, if we now drop the asterisks the problem discussed in Sec. 3 is reproduced with

$$g = \frac{d\zeta_0}{dy}. \tag{31}$$

The physical interpretation of the quantity τ defined by (12) is also obvious at once. It is $(1/2\pi)$ times the total transfer of angular momentum per unit height across the latitude circle of radius r.

The disturbance equation for the spherical surface has been obtained and briefly discussed by Kuo.

Most of the results obtained in Secs. 3, 4 are applicable to the present case. There are, however, also essential differences. In the first place, the final solution has a meaning only when α is an integer. During the investigation of the eigenvalue problem, however, we shall of course investigate all positive values. Another essential modification is the following. Since $g \neq -w''$, the existence proof of Sec. 4 cannot be readily extended to the present case. It is also clear that the symmetry or asymmetry of the velocity distribution is not an essential feature. Rather the equality or inequality of the values of w at the roots of $g = 0$ is what distinguishes cases (a) and (b) in Sec. 4.

The details of the investigation are also modified because the velocity distribution is more complicated. Typical diagrams of $w(y)$ and $g(y)$ are shown in Fig. 5. It is clear that one may consider only half of the flow (say right), and that the boundary conditions are

$$\varphi = 0, \qquad \text{or} \qquad \varphi' = 0, \qquad \text{at} \qquad y = 0, \tag{32}$$

and

$$\varphi \to 0, \qquad \varphi'/\varphi = -\alpha \text{ in the limit as } y \to \infty. \tag{33}$$

But since these conditions give $\tau = 0$ at both ends, no essential modification is necessary. The diagrams for the distribution of τ are similar to those shown in Sec. 4. In particular, one notes that the angular momentum is transferred northwards by a neutral disturbance if $w_s < w_t$ (opposite to the case shown in Fig. 5), i.e., if the mean zonal wind at the point of minimum vorticity is lower than that at maximum vorticity. Otherwise, a damped disturbance is necessary for northward transfer of angular momentum.

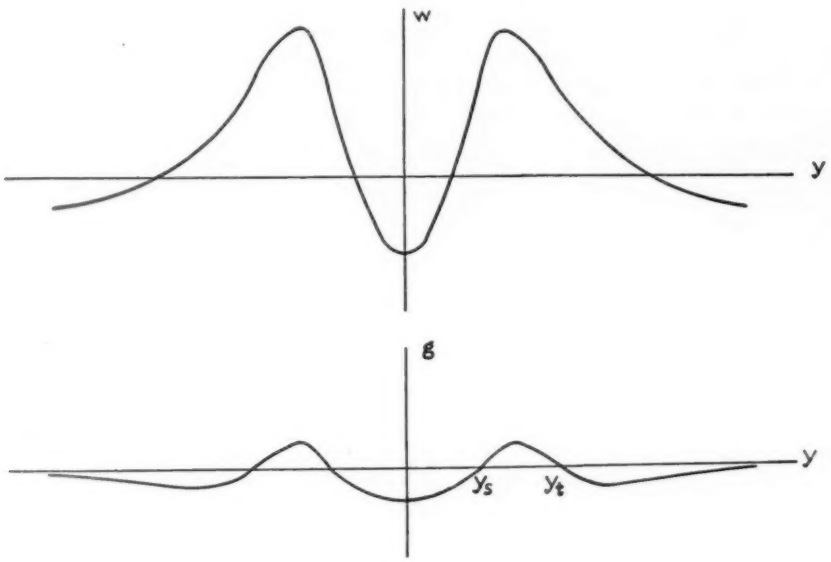

FIG. 5.

The existence of self-excited and neutral disturbances will now be discussed for the typical velocity and vorticity distribution shown in Fig. 5. We shall consider essentially the neutral disturbance in the case $w_s = w_t$, because the other cases can be covered by arguments similar to those used in Sec. 4. For this case, with $c = w_s = w_t$,

$$K \equiv \frac{g}{w - c}$$

is a regular non-negative function which approaches zero as $y \to \pm\infty$. Consider only $y > 0$. The equation

$$\varphi'' + \lambda\varphi + K\varphi = 0,$$

obtained from (9a) with $-\alpha^2$ replaced by λ, must now be shown to have a negative eigenvalue. Now, the first eigenvalue is [9]

$$\lambda_1 = \text{min.} \frac{\int_0^\infty (f'^2 - Kf^2)\, dy}{\int_0^\infty f^2\, dy}$$

for all functions f having proper differentiability properties, and satisfying the boundary condition $f'(0) = 0$ and making $\int_0^\infty f^2\,dy$ convergent. We want to show that for certain functions satisfying the boundary conditions the integral in the numerator is negative, for then λ_1 is negative and our result is proved.* For this purpose, take $f = e^{-\epsilon y^2}$. Then

$$\int_0^\infty (f'^2 - Kf^2)\,dy = \int_0^\infty (4\epsilon^2 y^2 - K)e^{-2\epsilon y^2}\,dy.$$

As $\epsilon \to 0$, this integral approaches $-\int_0^\infty K\,dy$. Hence, by taking some small but finite value for ϵ, the integral can be made negative, completing our proof.

6. **Validity of the asymptotic expansions.** We shall now slightly extend some theorems of Wasow on the asymptotic solutions of the equation of hydrodynamic stability of two-dimensional parallel flows. Wasow established the validity of the formal expansions in the neighborhood of a single critical point. We shall note the modifications necessary to make the theorems of Wasow applicable to cases with two critical points. The preceding sections have used the mathematical justification so derived.

We shall retain the notation of Wasow in the main, but with x_0 , x_0^* for the two isolated zeros; A_k , A_k^*, $k = 1, 2, 3$, for the exponential-type solutions; and so on (see Fig. 6).

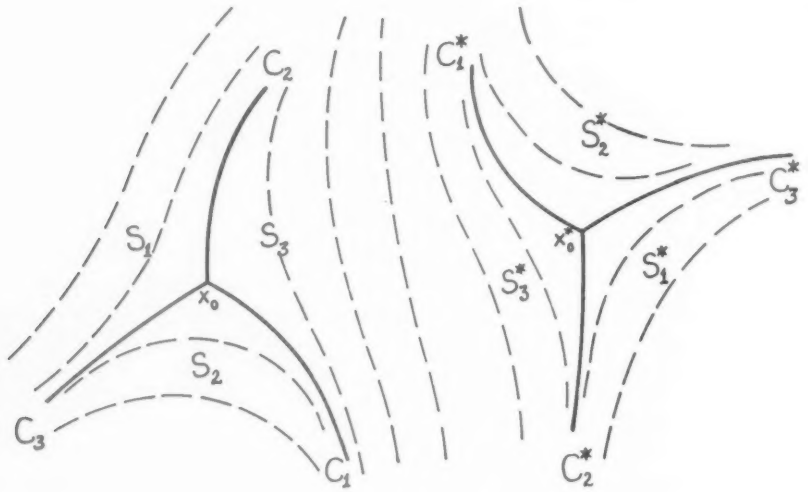

FIG. 6.

The jet flow profiles require the assumption of two zeros for the function $b_0(x)$ in the complex x-plane and further require that the region of validity of the asymptotic expressions have indefinite extent in order that the real x-axis may be used for certain integrals.

The following lemma is proved by Wasow, p. 857:

*This proof is similar to the one used by Friedrichs [6]; but the function $f(y)$ has to be chosen differently here.

LEMMA. *Let T^* be a subdomain of T every point of which can be reached from a fixed point p by a path Γ in T along which the τn quantities*

$$Re(P_i - P_j), \qquad (i = 1, 2, \cdots, \tau; j = 1, 2, \cdots, n),$$

are monotonically non-decreasing. Then there exist τ linearly independent solutions $y = y_r(x, \lambda)$, $r = (1, 2, \cdots, \tau)$ of (12) for which

$$y_r(x, \lambda) = e^{P_r(x,\lambda)}\left[\sum_{\nu=0}^{\nu_2} S_{r\nu}(x)\lambda^{-\nu/k} + E(T^*)\lambda^{-(\nu_2+1)/k}\right],$$

where ν_2 is an integer depending on the choice of ν_1, and $\lim_{\nu_1\to\infty} \nu_2 = \infty$.

Apparently to obtain a *definite* region of validity for the existence proof, Wasow identifies T with the doubly connected domain S, and T^* with $S - C_k$. However, the only special property required of the domain of validity, whatever its shape may be, is the monotonicity condition of the Lemma. By examining the lines $Re[\lambda Q(x)] = \text{const.}$ and $Re[\lambda Q^*(x)] = \text{const.}$ we can easily identify T and T^* with regions suitable for our purposes.

In view of the equation

$$Q(x) = Q(x_0^*) + Q^*(x), \tag{34}$$

it is easy to show that the necessary and sufficient condition that no C_k intersects any C_k^* is

$$Re[\lambda Q(x_0^*)] \neq 0. \tag{35}$$

This condition will be assumed since it holds for any reasonable velocity profile. Then one sector at x_0 has a common part with one sector at x_0^*, and by (34) the pair of curves bounding this domain and starting from x_0^* belong to the family

$$Re[\lambda Q(x)] = \text{constant}, \tag{36}$$

and indeed the constant is $\pm Re[\lambda Q(x_0^*)]$. Similarly the other pair of bounding curves belongs to the family

$$Re[\lambda Q^*(x)] = \text{constant}. \tag{37}$$

A set of contour lines (36) and (37) then fills the x-plane (Fig. 6) in an entirely regular manner except at x_0 and x_0^*; i.e., through each point is one and only one curve and any finite region containing neither x_0 nor x_0^* nor points of more than two sectors at x_0 and two sectors at x_0^* is mapped one-to-one to the corresponding region of the $\lambda Q(x)$ plane. For example, regions S_2, S_3, S_2^*, S_3^* are mapped as shown in Fig. 7. In this case the boundaries C_3, C_2 map into the cut C_0 through the origin, and C_2^*, C_3^* map into C_0^*. Any path Γ of importance here passes through the common sector S_3, S_3^* and we may take it to pass also into S_2 and S_2^* for definiteness. Then if we have chosen $Re[\lambda Q(x)] < 0$ in S_2, on crossing C_1 we have $Re[\lambda Q(x)] > 0$ on the remainder of Γ. This holds also if S_1^* is used in place of S_2^*. It is clear from the geometrical configuration that any two points in the plane can be joined by a path Γ on which $Re[\lambda Q(x)]$ is monotonically non-decreasing, where we select the subdominant determination of the real part of $\lambda Q(x)$ at the point p of the Lemma. This is essentially what Wasow does in choosing the point p_k on the boundary arc B_k of the sector S_k where $Re[\lambda Q(x)]$ takes on its minimum.

The coefficients of the basic differential equation may be taken to be such that the coefficients in the series (10), page 857 of [4], are regular analytic in the entire plane so that T can be identified with the plane, while T^* can be identified with any one of an infinite set of finite regions satisfying both the monotonicity property and the property

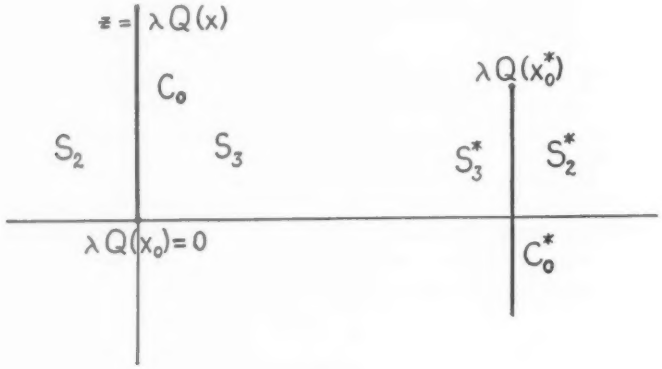

FIG. 7.

that the modulus of x can be taken indefinitely large. This is sufficient for the asymptotic expansions since x will be fixed at some *finite* point while $|\lambda| \to \infty$.

Then the theorems of Wasow follow formally by omitting the domain S from consideration and replacing it by any finite region of the above type.* In Sec. 2 we used this result with the following change of notation: x replaced by y; $-b_0(x)$ by $w(y) - c$; and $\{V, U_1, A_3, A_2\}$ by $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$ respectively.

*A different proof was given in the thesis, using only the final theorems of Wasow.

REFERENCES

1. C. C. Lin, *On the Stability of Two-Dimensional Parallel Flows*, Parts I, II, III; Q. Appl. Math. **3**, 117-142, 218-234, 277-301 (1945).
2. M. Lessen, *On the Stability of the Free Laminar Boundary Layer between Parallel Streams*, Thesis for Sc.D. at Mass. Inst. of Tech. (1948); also NACA Tech. Note No. 1929 (1947).
3. P. Chiarulli, *Stability of Two-dimensional Velocity Distributions of the Half-Jet Type*, Thesis for Ph.D. at Brown Univ. (1949); also Headquarters Air Material Command Technical Report No. F-TS-1228-IA (1949).
4. W. Wasow, *The Complex Asymptotic Theory of a Fourth Order Differential Equation of Hydrodynamics*, Annals Math., **49**, 852-871 (1948).
5. W. Tollmien, *Ein allgemeines Kriterium der Instabilitaet laminarer Geschwindigkeitsverteilungen*, Nachr. Ges. Wiss. Gottingen (neue Folge), **1**, 79-114 (1935).
6. R. von Mises and K. O. Friedrichs, *Fluid Dynamics*, Brown University Notes, pp. 205-207 (1941).
7. C. G. Rossby, *On a Mechanism for the Release of Potential Energy in the Atmosphere*, Jour. of Meteorology, **6**, 163-180 (1949).
8. Hsiao-lan Kuo, *Dynamic Instability of Two-Dimensional Nondivergent Flow in a Barotropic Atmosphere*, Jour. of Meteorology, **6**, 105-122 (1949).
9. R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, vol. 1, p. 344 ff.

THE REFLECTION OF AN ELECTROMAGNETIC PLANE WAVE BY AN INFINITE SET OF PLATES III*

BY

ALBERT E. HEINS

Carnegie Institute of Technology

1. Introduction. In 1947, J. F. Carlson and the present writer [1], [2] investigated the reflection properties of an electromagnetic plane wave by an infinite set of parallel plates. The problem was investigated for two polarizations, that is, the cases for which (a) the electric vector was parallel to the edges of the parallel plates and (b) the magnetic vector was parallel to the edges of the plates. Let us remark in passing, that case (b) is equivalent to a problem in acoustics. One may see this very easily by referring to CH II and noting that the problem solved there is a scalar electromagnetic problem and one can pair electromagnetic terminology with acoustic terminology and note the equivalence.

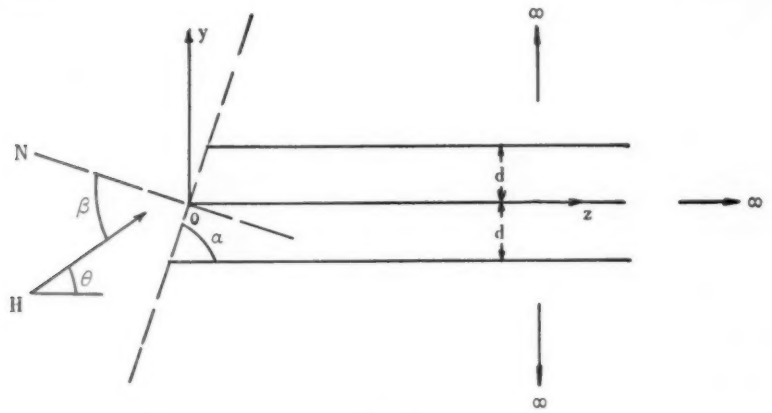

FIG. 1

In CH I and CH II, we solved the problems subject to the fact that there be a single propagating mode in the ducts and a single reflected wave in free space. We shall still retain the first hypothesis but shall reinvestigate the second one. We shall be concerned with the case in which there are two reflected waves and shall find the restrictions on wave length, angle of stagger and angle of incidence. We shall also find the new transmission coefficient as well as the two reflection coefficients. As a check, it will be shown that there exists a relation between the magnitudes of these coefficients which can be calculated with the aid of the complex form of the Poynting vector theorem. The case that we consider here is the "E plane" case, the one which has the acoustical analogue.

Since the formulation of this problem has been discussed in CH I and CH II, we only summarize the pertinent equations we used in CH II. We refer to Figure 1 for a view of the structure in any plane parallel to the $y\ z$ plane.

*Received Oct. 19, 1949. The material treated here is taken from an invited paper presented to the U.R.S.I.–I.R.E. meeting in Washington, D. C., on October 31, 1949.

The angle of incidence of the plane wave is θ; the angle of stagger of the parallel plates, which are semi-infinite in the z direction and infinite in the x direction, is α. The line $O\,N$ is perpendicular to the trace of the edges and the angle β [3] is the angle which the direction of propagation of the plane wave makes with this normal. Furthermore we have that $\alpha - \pi < \theta < \alpha$ and $0 < \alpha \leq \pi/2$. It is assumed here that the incident wave has only one component of the magnetic field, $H_x(y, z)$, that is, the one which is parallel to the x axis. For such an excitation, no other components of the magnetic field can be excited and hence the electric field may be derived from $H_x(y, z)$. If we now assume that the time variation of the field quantities is monochromatic, that is of the form $\exp(-ikct)$, then the Maxwell equations assume the form

$$\nabla \times \mathbf{E} = ik\mathbf{H} \tag{1.1}$$

and

$$\nabla \times \mathbf{H} = -ik\mathbf{E} \tag{1.2}$$

where $k = 2\pi/\lambda$ and λ is the free space wave-length. Here c is the velocity of light in units appropriate to those of λ. Let us, as a matter of notation, introduce $H_x(y, z) = \psi(y, z)$. Then (1.1) and (1.2) reduce to

$$\frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi = 0 \tag{1.3}$$

and this is to be solved subject to the boundary condition that the z component of the electric field vanishes on $y = nd$, $z \geq nd \cot \alpha$. If this boundary condition is expressed in terms of $\psi(y, z)$ we see that this is equivalent to the vanishing of $\partial\psi/\partial y$ on $y = nd$, $z \geq nd \cot \alpha$, $n = 0, \pm 1, \cdots$. There are also conditions at infinity which have been described in CH I.

We recall from CH II, Eq. (3.2) that this boundary value problem may be formulated as a Wiener-Hopf integral equation for the current density on the plates. The plates are assumed to be perfect conductors of zero thickness. That is

$$\begin{aligned} 0 = {} & ik \sin\theta \exp[ikz \cos\theta] \\ & + \frac{i}{4} \sum_{p=-\infty}^{\infty} \int_0^{\infty} I_0(z') \exp[ik\mu p] \frac{\partial^2}{\partial y^2} H_0^{(1)}[k\{(z - z' + pg)^2 + (y - nd - pd)^2\}^{1/2}]\, dz' \end{aligned} \tag{1.4}$$

where $z \geq 0$. Here $\mu = g \cos\theta + d \sin\theta$, and $g = d \cot \alpha$, while the variable y is set equal to nd. Further the surface current density on the m plate is proportional to $I_m(z)$ and may be expressed in terms of a single current density by the relation

$$I_m(z - mg) = I_0(z) \exp[ikm\mu].$$

For all y and z we have then the following Fourier integral representation for $\psi(y, z)$.

$$\begin{aligned} \psi(y, z) = {} & \exp[ik(y \sin\theta + z \cos\theta)] - \frac{1}{4\pi} \int_C dw\, J(w) \exp[i(wz + k\mu n - wgn)] \\ & \times \frac{\cos\{(y - nd - d)(k^2 - w^2)^{1/2}\} - \exp[i(k\mu - wg)] \cos\{(dn - y)(k^2 - w^2)^{1/2}\}}{\cos[d(k^2 - w^2)^{1/2}] - \cos(k\mu - wg)} \end{aligned} \tag{1.5}$$

where $J(w)$ is

$$J(w) = \int_0^\infty I_0(z') \exp(-iwz')\, dz'.$$

The path C is drawn within a certain strip of the complex w plane which we shall define in the following section. Here n is the largest integer in the ratio y/d. We assume here that k has a small positive imaginary part, and the precise meaning of small will be defined in Sec. 2.

2. The Fourier transform of the kernel of equation (1.4). The Fourier Transform of the kernel of the integral equation (1.4) is

$$-\frac{i}{2}\{k^2 - w^2\}^{1/2} \sum_{p=-\infty}^{\infty} \exp[ip(k\mu - wg) + |p| d\{k^2 - w^2\}^{1/2}]. \tag{2.1}$$

We choose here that branch of $(k^2 - w^2)^{1/2}$ which reduces to k when $w = 0$. The series in (2.1) is not in a particularly useful form since it does not exhibit the singularities of the function it represents, save for the possible branch points at $w = \pm k$. We say possible branch points advisedly, since in closing this sum, we shall find that they disappear and the only singularities possessed by (2.1) are two infinite sequences of poles. In any event, each term of the series has branch points at $w = \pm k$, so that each is regular in the strip $-\Im m k < \Im m w < \Im m k$. We observe further that the series (2.1) diverges when

$$k\mu - wg + d\{k^2 - w^2\}^{1/2} = 2n\pi \tag{2.2a}$$

or

$$k\mu - wg - d\{k^2 - w^2\}^{1/2} = 2n\pi \tag{2.2b}$$

$n = 0, \pm 1, \pm 2, \cdots$. From Eqs. (2.2a) and (2.2b) it is clear that the singularities of (2.1) are intimately related to the values of α, θ, d and k which we use. For $n = 0$, we find that w may be $k \cos(2\alpha - \theta) = \sigma_2$ or $k \cos\theta = \sigma_1$. Since $\alpha - \pi < \theta < \alpha$ and $0 < \alpha \leq \pi/2$, it is clear that $\Im m \sigma_2 < \Im m \sigma_1$. From this we see that the series converges at best in a strip $\Im m \sigma_2 < \Im m w < \Im m \sigma_1$ and this is so, provided that there are no other singularities given by Eqs. (2.2a) and (2.2b) which are in a smaller strip.

A next possible root which may be in the strip $\Im m \sigma_2 < \Im m w < \Im m \sigma_1$ comes from Eqs. (2.2a) and (2.2b) by putting $n = \pm 1$. Let us assume that the angle $\beta = \pi/2 + \theta - \alpha$ satisfies the inequality $0 < \beta \leq \pi/2$. Upon solving Eqs. (2.2a) and (2.2b) for $n = 1$ we get

$$w_1 = (2\pi/d) \sin\alpha[-\cos\alpha(1 - k\mu/2\pi) + \Delta] \tag{2.3a}$$

and

$$w_2 = (2\pi/d) \sin\alpha[-\cos\alpha(1 - k\mu/2\pi) - \Delta] \tag{2.3b}$$

where Δ is given by the expression

$$\Delta = [(dk/2\pi)^2 - \sin^2\alpha(1 - k\mu/2\pi)^2]^{1/2}$$

and the positive root is understood for Δ. The roots w_1 and w_2 are written in this form under the assumption that Δ is real for k real. Observe that w_1 satisfies Eq. (2.2a) while w_2 satisfies Eq. (2.2b) for $n = 1$.

We now write $\Delta_r = \Delta$ for k real and assume further that Δ does not vanish. Then we may expand w_1 and w_2 for $\Im m k = \epsilon$ sufficiently small (where now ϵ is small enough to insure the convergence of the expansion) and we obtain

$$\Im m w_1 = [(\epsilon \sin \alpha)/d][\mu \cos \alpha + (d^2 k_r + 2\pi\mu \sin^2 \alpha - k_r \mu^2 \sin^2 \alpha)/2\pi\Delta_r]$$

and

$$\Im m w_2 = [(\epsilon \sin \alpha)/d][\mu \cos \alpha - (d^2 k_r + 2\pi\mu \sin^2 \alpha - k_r \mu^2 \sin^2 \alpha)/2\pi\Delta_r]$$

if we only retain first powers of ϵ. In terms of the angle β, these expressions may be simplified to read

$$\Im m w_1 = \epsilon(\cos \alpha \sin \beta + \sin \alpha(dk_r \cos^2 \beta + 2\pi \sin \alpha \sin \beta)/2\pi\Delta_r]$$

and

$$\Im m w_2 = \epsilon[\cos \alpha \sin \beta - \sin \alpha(dk_r \cos^2 \beta + 2\pi \sin \alpha \sin \beta)/2\pi\Delta_r].$$

Clearly, under the restrictions we have imposed on α and β, $\Im m w_2 < \Im m w_1$ so that it remains to compare the relative orders of magnitude of $\Im m w_1$, $\Im m w_2$, $\Im m \sigma_1$ and $\Im m \sigma_2$. We shall now show that $\Im m w_2 < \Im m \sigma_1$. This implies that

$$\cos \alpha \sin \beta - \sin \alpha(dk_r \cos^2 \beta + 2\pi \sin \alpha \sin \beta)/2\pi\Delta_r < \cos \theta$$

or

$$\cos \alpha \sin \beta - \sin \alpha(dk_r \cos^2 \beta + 2\pi \sin \alpha \sin \beta)/2\pi\Delta_r < \sin (\alpha + \beta)$$

From this, we get immediately

$$-\sin \alpha(dk_r \cos^2 \beta + 2\pi \sin \alpha \sin \beta)/2\pi\Delta_r < \sin \alpha \cos \beta.$$

Since the left side of this inequality is always negative for α and β positive and acute, while the right side is positive, it is clear that $\Im m w_2 < \Im m \sigma_1$. We have similarly that $\Im m \sigma_2 < \Im m w_1$, $\Im m \sigma_1 < \Im m w_1$, and $\Im m w_2 < \Im m \sigma_2$. If we now combine all of the inequalities we get

$$\Im m w_2 < \Im m \sigma_2 < \Im m w < \Im m \sigma_1 < \Im m w_1$$

Hence the series (2.1) converges in the strip

$$\Im m \sigma_2 < \Im m w < \Im m \sigma_1$$

provided now that there are no other roots of Eqs. (2.2a) or (2.2b) in this strip. Such will be the case, if we limit dk_r properly. Let us suppose for the present that such is the case.

One observes at this point, that if Δ_r is real for real k, that is for $k = k_r$, then the following inequality is satisfied by k_r, d, θ and α. We have here

$$\frac{dk_r}{2\pi} > \frac{\sin \alpha}{1 + \cos (\theta - \alpha)}$$

But in order that one mode propagate in the ducts,

$$0 < (dk_r)/2\pi < \tfrac{1}{2}$$

Hence, subject to the condition that $0 < \beta < \pi/2$, the strip of regularity of (2.1) is as we have given it, provided

$$0 < \frac{\sin \alpha}{1 + \cos(\theta - \alpha)} < \frac{dk_r}{2\pi} < \frac{1}{2}.$$

This inequality on $(dk_r)/2\pi$ assures us of the existence of at least the two propagating reflected waves in free space. Since we desire only two such waves, we should put down conditions which insure that the remaining Δ_n for $n = -1, \pm 2, \pm 3, \cdots$ be imaginary for k_r real. In the first place, the remaining roots of Eqs. (2.2a) and (2.2b) are

$$[(2\pi n \sin \alpha)/d][-\cos \alpha(1 - k\mu/2\pi n) \pm i\Delta_n]$$

where

$$\Delta_n = [\sin^2 \alpha(1 - k\mu/2\pi n)^2 - (dk)^2/4n^2\pi^2]^{1/2}.$$

Δ_n will vanish if

$$(dk)/2n\pi = -\sin \alpha/[1 - \cos(\theta - \alpha)] \qquad n = \pm 1, \pm 2, \cdots \tag{2.4}$$

or

$$(dk)/2n\pi = \sin \alpha/[1 + \cos(\theta - \alpha)] \qquad n = \pm 1, \pm 2, \cdots. \tag{2.5}$$

But for n negative only, the roots

$$dk/2n\pi = -\sin \alpha/[1 - \cos(\theta - \alpha)] \qquad n = -1, -2, \cdots \tag{2.6}$$

satisfy the inequality $dk/2\pi > 0$, while for n positive, only the roots

$$dk/2n\pi = \sin \alpha/[1 + \cos(\theta - \alpha)] \qquad n = 1, 2, \cdots \tag{2.7}$$

satisfy the same inequality.

From the two sequence of roots (2.6) and (2.7) we can now form the condition that there be only two reflected waves. One reflected wave arises from the root $w = \sigma_2$. This we have seen in CH II corresponds to the reflected wave which makes an angle with respect to ON equal numerically to that of the incident wave. The other one corresponds to $w = w_2$. In order that a third reflected wave not be present

$$(dk_r)/2\pi < (2 \sin \alpha)/[1 + \cos(\theta - \alpha)] \qquad \text{or} \qquad \sin \alpha/[1 - \cos(\theta - \alpha)].$$

If dk_r satisfies this inequality, it satisfies the inequality

$$(dk_r)/2\pi < (n \sin \alpha)/[1 + \cos(\theta - \alpha)] \qquad \begin{matrix} 0 < \alpha - \theta < \pi/2 \\ n = 3, 4, \cdots \end{matrix}$$

as well as

$$(dk_r)/2\pi < (n \sin \alpha)/[1 - \cos(\theta - \alpha)] \qquad \begin{matrix} 0 < \alpha - \theta < \pi/2 \\ n = 2, 3, \cdots. \end{matrix}$$

We have then the following condition for two reflected waves.

$$0 < (\sin \alpha)/[1 + \cos(\theta - \alpha)] < (dk_r)/2\pi < \{\tfrac{1}{2} \text{ or } (2 \sin \alpha)/[1 + \cos(\theta - \alpha)]$$

$$\text{or } \sin \alpha/[1 - \cos(\theta - \alpha)]\}$$

The term on the right side is to be interpreted as the smaller of the three terms. For example if $\sin \beta = .25$, this inequality reads

$$0 < (4 \sin \alpha)/5 < (dk_r)/2\pi < \{\tfrac{1}{2} \text{ or } (4 \sin \alpha)/3\}$$

On the other hand, if $\sin \beta = .5$, this inequality reads

$$0 < (2 \sin \alpha)/3 < (dk_r)/2\pi < \{\tfrac{1}{2} \text{ or } (4 \sin \alpha)/3\}.$$

It is clear that we can find conditions for three or more reflected waves by arguments similar to those that we have just used, but we shall not pursue this detail any further. One final remark is in order. Suppose that $\pi/2 < \alpha - \theta < \pi$. In this case the inequality reads

$$0 < (\sin \alpha)/[1 - \cos (\theta - \alpha)] < (dk_r)/2\pi < \{\tfrac{1}{2} \text{ or } (2 \sin \alpha)/[1 - \cos (\theta - \alpha)]$$

$$\text{or } \sin \alpha/[1 + \cos (\theta - \alpha)]\}$$

with the same interpretation lent to the right side of the inequality as we did for the case $0 < \alpha - \theta < \pi/2$.

3. The solution of equation (1.4). We define Eq. (1.4) for all z to read

$$F(z) = \psi_0(z) + \frac{i}{4} \sum_{p=-\infty}^{\infty} \int_0^{\infty} I_0(z') \exp [ik\mu p] \frac{\partial^2}{\partial y^2} H_0^{(1)}[k\{(z - z' + pg)^2 + (y - nd - d)^2\}^{1/2}] \, dz' \tag{3.1}$$

where $F(z)$ vanishes identically for positive z. Further $\psi_0(z) \equiv ik \sin \theta \exp [ikz \cos \theta]$ for positive z and vanishes identically for negative z. In view of the restrictions we have placed on dk_r, θ and α, we have that

$$F(z) = O[\exp (i\sigma_2 z)] + O[\exp (iw_2 z)] \tag{3.2}$$

for $z \to -\infty$. In order to see this, we rewrite Eq. (3.1) in Fourier integral representation as

$$F(z) = \psi_0(z) - \frac{1}{4\pi} \int_C dw \, J(w) \exp [i(wz + k\mu n - wgn)] \cdot \frac{\{k^2 - w^2\}^{1/2} \sin [d\{k^2 - w^2\}^{1/2}]}{\cos [d\{k^2 - w^2\}^{1/2}] - \cos (k\mu - wg)}.$$

Since $I_0(z)$ has the same z dependence as the magnetic field, and since only the lowest mode propagates in the ducts, $J(w)$ is regular in the lower half plane $\Im m w < \Im m k$. Further since there are two reflected waves, the path C is drawn in the strip $\Im m \sigma_2 < \Im m w < \Im m \sigma_1$[4]. In order to find $F(z)$ for $z \to -\infty$, we need only calculate the residues of the poles in the lower half plane $\Im m w < \Im m \sigma_1$. Upon doing so, we find that the dominant terms are of the form given in Eq. (3.2); all other terms which have been omitted are small compared to these two. This incidently tells us that the unilateral Fourier transform of $F(z)$ is regular in the upper half plane $\Im m w > \Im m \sigma_2$, provided we assume the integrability of $F(z)$ for z negative and finite. This may be verified with the solution of the integral equation (3.1). The Fourier transform of $\psi_0(z)$ is regular in the lower

half plane $\Im mw < \Im m\sigma_1$. Finally the transform of the kernel (2.1) may be written in closed form as

$$\frac{\{k^2 - w^2\}^{1/2} \sin [d\{k^2 - w^2\}^{1/2}]}{\cos [d\{k^2 - w^2\}^{1/2}] - \cos [k\mu - wg]} = \frac{K_-(w)}{K_+(w)}$$

and this is regular in the strip $\Im m\sigma_2 < \Im mw < \Im m\sigma_1$. There is then a common strip of analyticity to all transforms involved and it is thus permissable to apply the Fourier transform to Eq. (3.1) to get

$$D(w) = \frac{k \sin \theta}{w - \sigma_1} - \frac{K_-(w)J(w)}{2K_+(w)} .$$

If we separate terms as we did in CH II we get

$$J(w) = \frac{2k \sin \theta K_+(\sigma_1)}{(w - \sigma_1)K_-(w)}$$

and

$$D(w) = \frac{k \sin \theta[K_+(w) - K_+(\sigma_1)]}{(w - \sigma_1)K_+(w)}$$

The integral function of separation vanishes identically in this case. The terms $K_-(w)$ and $K_+(w)$ are now given by the following expressions.

$$K_-(w) = N_-(w)/D_-(w) \qquad \text{and} \qquad K_+(w) = D_+(w)/N_+(w)$$

where

$$N_-(w) = \{\exp [x(w)]\{k - w\} \prod_{n=1}^{\infty} [\{1 - (kd/n\pi)^2\}^{1/2} + iwd/n\pi] \exp [-iwd/n\pi]$$

$$N_+(w) = \{2d(k + w)$$

$$\times \exp [-\chi(w)]/(d^2 + g^2)\} \prod_{n=1}^{\infty} [\{1 - (kd/n\pi)^2\}^{1/2} - iwd/n\pi] \exp [iwd/n\pi]$$

$$D_-(w) = (w - \sigma_1)\left\{ \prod_{n=-\infty}^{-1} [\Delta_n - i\Psi_n] \exp [(k\mu - wg + iwd)/2n\pi + i(\pi/2 - \alpha)]\right\}$$

$$\times \left\{ \prod_{n=2}^{\infty} [\Delta_n + i\Psi_n] \exp [(k\mu - wg - iwd)/2n\pi - i(\pi/2 - \alpha)]\right\}$$

$$\times \{(\Psi_1 - \Delta) \exp [(k\mu - wg + iwd)/2\pi - i(\pi/2 - \alpha)]\}$$

and

$$D_+(w) = (w - \sigma_2)\left\{ \prod_{-\infty}^{-1} [\Delta_n + i\Psi_n] \exp [(k\mu - wg - iwd)/2n\pi - i(\pi/2 - \alpha)]\right\}$$

$$\times \left\{ \prod_{n=2}^{\infty} [\Delta_n - i\Psi_n] \exp [(k\mu - wg + iwd)/2n\pi + i(\pi/2 - \alpha)]\right\}$$

$$\times \{(\Psi_1 + \Delta) \exp [(k\mu - wg - iwd)/2\pi + i(\pi/2 - \alpha)]\}.$$

Here

$$\Delta_n = \{\sin^2 \alpha(1 - k\mu/2n\pi)^2 - (kd)^2/4n^2\pi^2\}^{1/2} \qquad n = -1, \pm 2, \cdots$$

$$\Delta = \{(kd)^2/4\pi^2 - \sin^2 \alpha(1 - k\mu/2\pi)^2\}^{1/2}$$

$$\Psi_n = [\cos \alpha(1 - k\mu/2n\pi)] + (wd \csc \alpha)/2n\pi \qquad n = \pm 1, \pm 2, \cdots$$

and

$$\chi(w) = -[iwd/2\pi][(\alpha - \pi/2) \cot \alpha - \ln (2 \sin \alpha)].$$

The subscripts $-$ and $+$ on K, N or D indicate that the function is regular [5] in the lower or upper half plane, as the case may be. For example $D_-(w)$ is regular in the lower half plane $\Im m w < \Im m \sigma_1$, $N_-(w)$ is regular in the lower half plane $\Im m w < \Im m k$, etc.

4. The investigation of the far fields. We have found $J(w)$, the Fourier transform of $I_0(z)$. If we now use Eq. (1.5), we have an integral representation for $\psi(y, z)$. From this representation we can determine $\psi(y, z)$ for $z \to \infty$ (in the ducts) and $\psi(y, z)$ for $z \to -\infty$, (in free space). For $z > 0$, we close the path C in the upper plane [6] and obtain

$$\psi(y, z) = \frac{i \sin \theta K_+(\sigma_1)}{2d(k - \sigma_1)K_+(k)} [1 - \exp (ik\mu - ikg)] \exp [i(kz + k\mu n - kgn)]$$

$$+ O[(\cos \pi y/d) \exp [-z\{\pi^2/d^2 - k^2\}^{1/2}]]$$

Hence for $z \to \infty$, $\psi(y, z)$ is asymptotic to

$$\frac{i \sin \theta K_+(\sigma_1)}{2d(k - \sigma_1)K_+(k)} [1 - \exp (ik\mu - ikg)] \exp [i(kz + k\mu n - kgn)]$$

and from this we can find the transmission coefficient in the duct. It is

$$T = \frac{i \sin \theta K_+(\sigma_1)}{2d(k - \sigma_1)K_+(k)} [1 - \exp (ik\mu - ikg)].$$

For k real, $| T |^2$ reduces to

$$| T |^2 = \left(\frac{\Psi_1 - \Delta}{\Psi_1 + \Delta}\right)_k \left(\frac{\Psi_1 + \Delta}{\Psi_1 - \Delta}\right)_{\sigma_1} \frac{\sin^2 (\alpha - \theta)}{[\cos^2 \theta/2][\sin^2 (\alpha - \theta/2)]}$$

where the subscripts k and σ_1 on the first two fractions indicate that w has been evaluated at these values.

For $z < 0$, we close the path C in the lower half plane and obtain

$$\psi(y, z) = \exp [ik(y \sin \theta + z \cos \theta)]$$

$$+ \frac{J(\sigma_2) \exp [ik\{y \sin (2\alpha - \theta) + z \cos (2\alpha - \theta)\}]}{2d[\cot (2\alpha - \theta) - \cot \alpha]}$$

$$- \frac{J(w_2)\{k^2 - w_2^2\}^{1/2} \exp [i\{w_2 z + (k^2 - w_2^2)^{1/2}\}]}{4\pi(\csc \alpha)\Delta} +$$

terms which attenuate exponentially for $z \to -\infty$. The term $(k^2 - w_2^2)^{1/2}$ may be written as $[2\pi/d]\,[\Delta \cos\alpha - \sin^2\alpha(1 - k\mu/2\pi)]$. Hence the first reflection coefficient is

$$R_1 = J(\sigma_2)/2d[\cot(2\alpha - \theta) - \cot\alpha]$$

while the second reflection coefficient is

$$R_2 = -J(w_2)(k^2 - w_2^2)^{1/2}/4\pi\Delta \csc\alpha.$$

For k real

$$|R_1|^2 = \frac{\sin^2\theta}{(\sigma_2 - \sigma_1)^2 \sin^2(2\alpha - \theta)} \left| \frac{K_+(\sigma_1)}{K_+'(\sigma_2)} \right|^2$$

where the prime in $K_+(w)$ denotes differentiation with respect to the argument w. Upon simplifying further, we get finally for k real

$$|R_1|^2 = [\cot^2(\alpha - \theta/2)][\tan^2\theta/2]\left[\frac{\Psi_1 - \Delta}{\Psi_1 + \Delta}\right]_{\sigma_1}\left[\frac{\Psi_1 + \Delta}{\Psi_1 - \Delta}\right]_{\sigma_1}.$$

Furthermore for k real

$$|R_2|^2 = \frac{4k^2[\sin^2\alpha][\tan^2\theta/2][\sin^2(\alpha - \theta)]}{[w_2 - \sigma_1][w_2 - \sigma_2]}\left[\frac{\Psi_1 + \Delta}{\Psi_1 - \Delta}\right]_{\sigma_1}\left[\frac{k + w_2}{k - w_2}\right].$$

The values which we have given for $|T|^2$, $|R_1|^2$ and $|R_2|^2$ have been derived subject to the conditions

(a) $$0 < \frac{\sin\alpha}{1 + \sin\beta} < \frac{dk}{2\pi} < \left\{\frac{1}{2} \text{ or } \frac{2\sin\alpha}{1 + \sin\beta} \text{ or } \frac{\sin\alpha}{1 - \sin\beta}\right\}$$

(b) $$0 < \beta < \pi/2.$$

The modification for $-\pi/2 < \beta < 0$ is quite direct. We observe that for this case, $\pi/2 < \alpha - \theta < \pi$. The requirement for two reflected waves for this interval of $\alpha - \theta$ is

(c) $$0 < \frac{\sin\alpha}{1 - \sin\beta} < \frac{kd}{2\pi} < \left\{\frac{1}{2} \text{ or } \frac{2\sin\alpha}{1 - \sin\beta} \text{ or } \frac{\sin\alpha}{1 + \sin\beta}\right\}.$$

In view of the requirement on $\alpha - \theta$, it is clear that the inequalities (a) and (c) are equivalent. There are a number of modifications in the expressions for $|R_1|^2$, $|R_2|^2$, and $|T|^2$ for the range $-\pi/2 < \beta < 0$, but we shall not pursue the matter further.

5. Conservation of power flow. As a check for our calculations of $|T|$, $|R_1|$ and $|R_2|$ we shall now see what information we can derive from the Poynting vector theorem. In the first place, the power flow from free space into any duct is (per unit length in the x direction)

$$\Re e \int_0^d \left(\frac{\csc\alpha}{2}\right)(\mathbf{E} \times \mathbf{H}^*)\cdot\mathbf{n}\, dy$$

where $\mathbf{n}$ is the normal vector into the parallel plate region. In terms of ψ, this may be written as

$$\Re e \frac{i \csc\alpha}{2} \int_0^d \left[\cos\alpha \frac{\partial\psi}{\partial y} - \sin\alpha \frac{\partial\psi}{\partial z}\right]\psi^*\, dy. \tag{4.1}$$

Here $y = z \tan \alpha$, that is we are integrating over the opening of a parallel plate region. Hence upon simplifying the parenthesis and integrating over y, we get

$$\frac{d \csc \alpha}{2} [k \sin (\alpha - \theta) - k \mid R_1 \mid^2 \sin (\alpha - \theta) - 2\pi\Delta \mid R_2 \mid^2/d].$$

This represents the average flow of power into a given duct per unit length in the x direction. This in turn is to be balanced by the average flow of power far down the duct, that is as $z \to \infty$, $nd < y < (n + 1)d$. The magnetic field in the duct for $z \to \infty$, $nd < y < (n + 1)d$ is

$$\psi(y_1 z) = T_n \exp [ikz].$$

Hence the power flow down any duct for $z \to \infty$ is $\mid T \mid^2 d/2$ (recall that $\mid T_n \mid$ is independent of n, the duct number). Hence conservation of power flow demands that

$$k \mid T \mid^2 = \csc \alpha [k \sin (\alpha - \theta) - k \mid R_1 \mid^2 \sin (\alpha - \theta) - 2\pi\Delta \mid R_2 \mid^2/d]. \tag{4.2}$$

Upon substituting in $\mid R_1 \mid^2$ and $\mid R_2 \mid^2$ we find that Eq. (4.2) is an identity, thus verifying the conservation theorem.

6. **Some concluding remarks.** If we try to find conditions for higher reflected waves, that is the cases corresponding to $n = -1, -2, \cdots$ or $n = 2, 3, \cdots$, we find that the conditions for a given number of reflected waves become fairly complex. On the other hand, a study of the zero of the Δ_n's and their ordering will show us how each reflected wave enters. These results will depend on α and β. The case $\beta = 0$, has been excluded from our present discussion, but it offers no difficulties. In this case, we get higher reflected waves in pairs. The result is not surprising since the incident wave normal and the normal to the trace of the edges of the plates are parallel.

For the case of a single reflected wave, we noted that the reflected wave has a normal whose direction is independent of k. In the case of two reflected waves, the first reflected wave has a normal whose direction is independent of k, but the second reflected wave has a direction which is dependent on k. We shall present a discussion of the numerical results elsewhere.

Finally, we observe that $\mid R_1 \mid^2$, $\mid R_2 \mid^2$ and $\mid T \mid^2$ have not been normalized. If we refer back to Eq. (4.2) and divide through by $k \csc \alpha \sin (\alpha - \theta)$, the normalized first and second reflection coefficients, as well as transmission coefficients for the case discussed, are respectively

$$\mid R_1 \mid^2, \qquad 2\pi\Delta \csc (\alpha - \theta) \mid R_2 \mid^2/d \qquad \text{and} \qquad \mid T \mid^2 \sin \alpha \csc (\alpha - \theta).$$

For the normalized coefficients, all magnitudes are numerically less than unity.

References and Notes.

1. J. F. Carlson and A. E. Heins *The reflection of an electromagnetic plane wave by an infinite set of plates, I*, this Quarterly, **4**, 313-329 (1947). Hereafter we shall refer to this paper as CH I.

2. A. E. Heins and J. F. Carlson *The reflection of an electromagnetic plane wave by an infinite set of plates, II*, this Quarterly, **5**, 82-88 (1947). Hereafter we shall refer to this paper as CH II.

3. If the propagation normal of the incident wave falls to the left of ON, angle β is a positive acute angle, while if it falls to the right of ON, the angle β is a negative acute angle. The case $\beta = 0$ requires separate treatment.

4. At this point we realize that to get more than two reflected waves, the inequalities in sec. 2 have to be modified. As we carry on this modification we find that an indefinite number of reflected waves cannot enter. For example, if $\alpha = \pi/2$, then the inequality for two reflected waves reads

$$0 < \frac{1}{1 + \sin\theta} < \frac{dk}{2\pi} < \left\{\frac{1}{2} \text{ or } \frac{2}{1 + \sin\theta} \text{ or } \frac{1}{1 - \sin\theta}\right\}$$

for $0 < \theta < \pi/2$. This is impossible since $\sin\theta < 1$. Hence there is only one reflected wave in this case. The formulation we gave in CH I makes no assumptions as to the form of $\psi(y, z)$ to the left of the parallel plates. The convergence study in sec 2, of the present paper gives us conditions for one, two, etc., reflected waves.

5. See CH I, p. 321 for some remarks regarding the present use of the term regular.

6. See CH I, sec. 4.

—NOTES—

A VARIATIONAL PROBLEM FOR THE ROOTS OF A CUBIC EQUATION*

(A contribution to the theory of servomechanisms)

By HANS BÜCKNER (*Minden, Germany*)

1. The problem. Let t denote the time, $u(t)$ a function of time and $v(t)$ the displacement of a servomechanism, which is to simulate $u(t)$. As regards the error

$$\eta(t) = v(t) - u(t)$$

many servomechanisms obey the equation

$$\eta''' + a_1\eta'' + a_2\eta' + a_3\eta = -u''' - a_4u'' \tag{1}$$

with constant coefficients $a_1 , \cdots a_4$. The quality of the servomechanism depends on the roots of the algebraic equation

$$x^3 + a_1x^2 + a_2x + a_3 = 0. \tag{2}$$

These may be denoted by x_1 , x_2 , x_3 . We consider the quantity

$$\xi = \text{Min } (-\Re x_1 , -\Re x_2 , -\Re x_3),$$

where $\Re x_i$ means the real part of x_i . Then $\xi > 0$ means stability, and the greater ξ, the faster the servomechanism follows sudden changes of $u(t)$. A positive ξ appears if and only if the well known Hurwitz-conditions are satisfied. These are

$$a_1 > 0, \qquad a_2 > 0, \qquad a_1a_2 > a_3 > 0.$$

If a_1 and a_2 are regarded as fixed, and a_3 as variable, ξ is a function of a_3 . An interesting problem both from a mathematical and a technical point of view, is to find Max ξ. There is a simple formula for it, which does not seem to be as well known as its merits.

2. The solution. Introducing the classic notations

$$p = a_2 - \frac{a_1^2}{3}, \qquad q = a_3 - \frac{a_1a_2}{3} + \frac{2}{27}a_1^3,$$

$$D = -4p^3 - 27q^2,$$

we may represent the roots by

$$\begin{aligned}
x_1 &= -\frac{a_1}{3} + \epsilon\left[-\frac{q}{2} + \frac{1}{18}(-3D)^{1/2}\right]^{1/3} + \epsilon^2\left[-\frac{q}{2} - \frac{1}{18}(-3D)^{1/2}\right]^{1/3}, \\
x_2 &= -\frac{a_1}{3} + \epsilon^2\left[-\frac{q}{2} + \frac{1}{18}(-3D)^{1/2}\right]^{1/3} + \epsilon\left[-\frac{q}{2} - \frac{1}{18}(-3D)^{1/2}\right]^{1/3}, \\
x_3 &= -\frac{a_1}{3} + \left[-\frac{q}{2} + \frac{1}{18}(-3D)^{1/2}\right]^{1/3} + \left[-\frac{q}{2} - \frac{1}{18}(-3D)^{1/2}\right]^{1/3},
\end{aligned} \tag{3}$$

*Received March 28, 1949.

where

$$\epsilon = -\frac{1}{2} + \frac{i}{2}(3)^{1/2}.$$

The quantity p depends on a_1 and a_2 only.

The roots x_1 , x_2 , x_3 are continuous functions of a_3 . Hence the same holds for ξ. The quantity ξ vanishes for $a_3 = 0$ and for $a_3 = a_1a_2$. In the first case one root vanishes, in the latter $x = \pm i(a_2)^{1/2}$ represents two roots with vanishing real parts.

For $0 < a_3 < a_1a_2$ we have $\xi > 0$. Therefore ξ as a function of a_3 has a positive maximum in the interval $(0, a_1a_2)$.

The criterion for multiple roots is $D = 0$. For $D > 0$ three different real roots exist; for $D < 0$ one pair of conjugate complex roots exists. The roots x_1 and x_2 of the system (3) give these complex roots, while x_3 is real.

The only case where D may vanish is for $p \leq 0$. Vanishing takes place for two values of a_3 which we denote by $(\alpha - \beta)$ and $(\alpha + \beta)$. We can write

$$\alpha = \frac{a_1a_2}{3} - \frac{2a_1^3}{27}, \qquad \beta = 2\left(-\frac{p}{27}^3\right)^{1/2}.$$

All roots are continuously differentiable functions of a_3 in all intervals which do not contain $(\alpha - \beta)$ or $(\alpha + \beta)$. Starting from the well-known formulae

$$x_1 + x_2 + x_3 = -a_1 ,$$

$$x_1^2 + x_2^2 + x_3^2 = a_1^2 - 2a_2 ,$$

$$x_1^3 + x_2^3 + x_3^3 = -a_1^3 + 3a_1a_2 - 3a_3 ,$$

we find

$$\frac{dx_1}{da_3} = \frac{1}{(x_1 - x_2)(x_1 - x_3)} ,$$

$$\frac{dx_2}{da_3} = \frac{1}{(x_2 - x_1)(x_2 - x_3)} , \tag{4}$$

$$\frac{dx_3}{da_3} = \frac{1}{(x_3 - x_1)(x_3 - x_2)} .$$

Now, let a_3 be a value for which $D > 0$. Then all roots are real and different from one another. Hence $-\xi$ coincides with one root, say x_i , for a sufficient small neighbourhood of a_3 . From (4) it follows that

$$\frac{d\xi}{da_3} = -\frac{dx_i}{da_3} \neq 0,$$

in this neighbourhood. Thus, a maximum of ξ can exist within the interval $(0, a_1a_2)$ only if $D \leq 0$.

Next, we consider a value a_3 and a neighbourhood of it in which $D < 0$. In this case we may write

$$\Re x_1 = \Re x_2 = -\frac{a_1}{3} - \frac{1}{2}\left[-\frac{q}{2} + \frac{1}{18}(-3D)^{1/2}\right]^{1/3} - \frac{1}{2}\left[-\frac{q}{2} - \frac{1}{18}(-3D)^{1/2}\right]^{1/3}$$

and, furthermore

$$\xi = \text{Min}\,(-\Re x_1\,,\ -x_3).$$

The criterion for

$$x_3 = \Re x_1$$

is $q = 0$. This means that $a_3 = \alpha$.

If it is assumed that a_3 does not coincide with α, ξ is represented either by $-x_3$ or by $-\Re x_i$ for a sufficient small neighbourhood of a_3. From (4) it follows that

$$\frac{d\xi}{da_3} = -\Re\frac{dx_i}{da_3} \neq 0$$

with a suitable subscript i. This holds for the neighbourhood. Therefore a maximum does not exist if $D < 0$ and $q \neq 0$.

Summing up, we may state that the maximum of ξ is given either by $D = 0$ or by $q = 0$, $D < 0$.

The following three cases may occur.

1) $p > 0$: We find $D < 0$. Max ξ is determined by $q = 0$. We have

$$\text{Max}\,\xi = \frac{a_1}{3}\,; \qquad a_3 = \frac{a_1a_2}{3} - \frac{2}{27}a_1^3\,.$$

2) $p = 0$: We find $D = 0$, $q = 0$ for $a_3 = \alpha$ and

$$\text{Max}\,\xi = \frac{a_1}{3} \qquad \text{for} \qquad a_3 = \frac{a_1a_2}{3} - \frac{2}{27}a_1^3\,.$$

3) $p < 0$: It is obvious that $D > 0$ for $q = 0$. The maximum of ξ is obtained for $D = 0$. The roots then are given by

$$-x_1 = -x_2 = \frac{a_1}{3} - \left(\frac{q}{2}\right)^{1/3}$$

$$-x_3 = \frac{a_1}{3} + 2\left(\frac{q}{2}\right)^{1/3}.$$

For $a_3 = \alpha \pm \beta$, we have $q = \pm\beta$ and

$$\xi = \frac{a_1}{3} - \left(\frac{\beta}{2}\right)^{1/3} \qquad \text{for} \qquad a_3 = \alpha + \beta,$$

$$\xi = \frac{a_1}{3} - 2\left(\frac{\beta}{2}\right)^{1/3} \qquad \text{for} \qquad a_3 = \alpha - \beta.$$

We see that $a_3 = \alpha + \beta$ gives the greater value for ξ. As the maximum of ξ exists, at least one of the abscissae $a_3 = \alpha \pm \beta$ must be in the interval $(0, a_1a_2)$. Since D is negative for $a_3 = a_1a_2$ (where complex roots exist), the abscissa $a_3 = \alpha + \beta$ surely is in the interval.

Therefore, we may state that

$$\text{Max}\,\xi = \frac{a_1}{3} - \frac{1}{(3)^{1/2}}(-p)^{1/2} \qquad \text{for} \qquad a_3 = \frac{a_1a_2}{3} - \frac{2}{27}a_1^3 + \frac{2}{9}(3)^{1/2}(-p)^{3/2}.$$

Obviously we can combine all cases and write

$$\text{Max } \xi = \Re\left\{\frac{a_1}{3} - \frac{1}{(3)^{1/2}}(-p)^{1/2}\right\}.$$

This is obtained for

$$a_3 = \Re\left\{\frac{a_1 a_2}{3} - \frac{2}{27}a_1^3 + \frac{2}{9}(3)^{1/2}(-p)^{3/2}\right\}.$$

PERIODIC MOTIONS OF A NON-LINEAR DYNAMIC SYSTEM*

By H. SERBIN (*Purdue University*)

1. Introduction. It is a matter of common knowledge that many problems of dynamics are non-linear in character and that the linearization of such problems is an expediency adopted in the face of mathematical difficulties attending the non-linearity. Fortunately, the linearized treatment gives useful results in a number of problems. On the other hand, there are cases where the essential characteristics of the phenomenon are altered by the assumption of linearity.

In the absence of general methods for handling non-linear problems, it is desirable to investigate typical non-linear phenomena to as great detail as possible with the intention of arriving at general methods of analysis.

One such problem is van der Pol's equation

$$x'' - \epsilon(1 - x^2)x' + x = 0, \qquad \left(x' = \frac{dx}{dt},\ x'' = \frac{d^2x}{dt^2}\right). \tag{1.1}$$

This defines the one-dimensional motion of a particle of unit mass suspended on a spring of unit stiffness and subjected to negative damping at low amplitudes and positive damping at high amplitudes.

The existence of periodic solutions of an equation of more general form

$$x'' + f(x)x' + g(x) = 0 \tag{1.2}$$

was considered by Liénard, Ref. 1, who proved, for $g(x) = x$ and under fairly general conditions on $f(x)$ (see Ref. 2), that there exists a periodic solution. The proof used in Ref. 1 was reproduced in Ref. 3 for $g(x) \neq x$.

In the present paper, the problem is considered from a different point of view. The periodicity of the solution is shown to exist for a wider class of differential equations (1.2) which can best be described and compared as in the following table:

	$F(x) = \int_0^x f(x)\,dx$	$G(x) = \int_0^x g(x)\,dx$	$f(x),\ g(x)$
Lefschetz	$F(\infty) = \infty$	$G(\infty) > 0$	Symmetric
Present paper	$F(\infty) > 0$	$G(\infty) = \infty$	No symmetry required
	$F(\infty) = \infty$	$G(\infty) > 0$	No symmetry required

*Received July 6, 1949.

In addition, the case in which $F(\infty)$ and $G(\infty)$ are positive and both finite is discussed. It is shown that an example can be constructed in which a periodic solution does not exist.

A number of inequalities relating to the initial energy and amplitude of motion are derived on the basis of which one can, in some cases, calculate upper and lower bounds of the amplitude of the periodic motion without solving the differential equation. This should be useful in application to practical problems. It is conceivable, in some instances, that an average of the upper and lower bounds will define the amplitude with sufficient accuracy for the purpose required.

Since the problem is non-linear, one must expect such inequalities to be non-linear, for instance, algebraic. The solution of these inequalities is carried out conveniently by graphical methods. This must not be compared with graphical methods of solution of the differential equation. It is conceivable that for a given degree of accuracy, examples may be constructed of the type (1.2) in which the graphical method of solution by isoclines would give results incorrect both qualitatively and quantitatively, showing a periodic solution where none exists.

It should be pointed out that the assumption of symmetry in $f(x)$ and $g(x)$ made heretofore is not consistent with the nature of the problem. If these functions are determined empirically, then one will not have symmetry in the mathematical sense if x and $-x$ represent two different positions of the system. One is accustomed, in small deflections, to replacing an empirical curve which is nearly parabolic at the origin by an exact parabola. However, for large deflections, this is not permissible. Any lack of symmetry vitiates the validity of the theorem of periodicity as known heretofore.

A mechanical system illustrating the non-symmetric case is afforded by a pendulum in which friction acts proportional to velocity. If one adds a device which acts through a limited range of motion of the pendulum at the bottom of the swing in such a way as to sense the velocity and then to force the motion proportional to the velocity over the friction, then one would have the features characteristic of Eq. (1.2). If the constant of proportionality is different for x and $-x$, the asymmetric case will exist.

2. Qualitative character of the solution. Consider the differential equation (1.2) and define

$$F(x) = \int_0^x f(x)\,dx,$$

$$G(x) = \int_0^x g(x)\,dx, \tag{2a}$$

$$H = \tfrac{1}{2}x^{\cdot 2} + G(x).$$

Make the assumptions:

(A) $f(x)$ and $g(x)$ are continuous for $x \geqq 0$,
(B) $f(x) < 0$ for $0 < x < x_1$,
(C) $f(x) > 0$ for $x > x_1$,
(D) $F(\infty) > 0$,
(E) $g(x) > 0$ for $x > 0$,
(F) $F(\infty)G(\infty) = \infty$.

Then $F(x)$ has only one positive zero, $x = x_2$.

Since

$$\frac{dH}{dt} = \dot{x}[\ddot{x} + g(x)],$$

it follows from Eq. (1.2) that

$$\frac{dH}{dt} = -f(x)\dot{x}^2, \tag{2.1}$$

which shows that H increases with time for $x < x_1$ and decreases for $x > x_1$. Dividing both sides of Eq. (2.1) by $\dot{x}$ and simplifying,

$$\frac{dH}{dx} = \mp f(x)[2(H - G)]^{1/2}, \tag{2.2}$$

where the sign is minus or plus according as the velocity is plus or minus.

Suppose H plotted versus x in an (x, H)-plane (Fig. 1). It is seen that $H \geqq G$, the equality implying that $dH/dx = 0$. At a point where $H > G$, the right side of Eq. (2.2) satisfies the Lipschitz condition in H. Since $f(x)$ and $G(x)$ are continuous, it follows from the fundamental existence theorem on the solution of ordinary differential equations that there is one and only one solution passing through (x, H) for $H > G$. Referring to Fig. 1, there are four distinct parts of the trajectory:

Part	$\dot{x}$	f	$\frac{dH}{dx}$
01	+	−	+
13	+	+	−
35	−	+	+
56	−	−	−

The subscripts 0, 1, etc., will indicate the value of a variable at the corresponding point of the trajectory.

It is clear that $H_1, H_2, \cdots, H_6, x_3$ are monotone, continuous, functions of H_0. The proof given in Ref. 3 shows, in addition, that $H_6 - H_0$ is a monotone increasing function of x_3.

The analysis in the following pages is directed toward showing that $H_6 - H_0$, regarded as a function of H_0 on the interval $0 \leqq H_0 < \infty$, changes sign. Because of the monotone character of $H_6 - H_0$, it will follow that there is one and only one trajectory for which $H_6 = H_0$.

3. Behaviour of the solution for $0 < x \leqq x_2$. Consider only those trajectories for which $x_3 \geqq x_2$. Define

$$I_{ij}(G) = \int_{ij} [2(H - G)]^{1/2}\, dH,$$

where the subscript ij indicates integration along the trajectory in the sense $i \to j$. If $G(x) < G'(x)$ on the part ij of the trajectory, then $I_{ij}(G) < I_{ij}(G')$. Since

$$I_{ij} = \mp \int_{ij} f(x)\,dx = \mp(F_j - F_i),$$

it follows that[1]

$$\begin{aligned} I_{01}(0) &= (2H_1)^{1/2} - (2H_0)^{1/2} < I_{01}(G) \\ &= -F_1 < I_{01}(G_1) = [2(H_1 - G_1)]^{1/2} - [2(H_0 - G_1)]^{1/2}, \end{aligned} \tag{3.1}$$

$$\begin{aligned} I_{21}(G_1) &= [2(H_1 - G_1)]^{1/2} - [2(H_2 - G_1)]^{1/2} < I_{21}(G) \\ &= -F_1 < I_{21}(G_2) = [2(H_1 - G_2)]^{1/2} - [2(H_2 - G_2)]^{1/2}. \end{aligned} \tag{3.2}$$

From (3.1), it follows that

$$\{[2H_1]^{1/2} + F_1\}^2 < 2H_0 < \{[2(H_1 - G_1)]^{1/2} + F_1\}^2 + 2G_1\,,$$

and from (3.2) that

$$\{[2(H_1 - G_1)]^{1/2} + F_1\}^2 + 2G_1 < 2H_2 < \{[2(H_1 - G_2)]^{1/2} + F_1\}^2 + 2G_2\,.$$

The latter two inequalities yield

$$\begin{aligned} 0 < 2(H_2 - H_0) &< -2F_1\{[2H_1]^{1/2} - [2(H_1 - G_2)]^{1/2}\} \\ &= \frac{-4F_1G_2}{\{[2H_1]^{1/2} + [2(H_1 - G_2)]^{1/2}\}} < \frac{-4F_1G_2}{(2H_0)^{1/2}}\,. \end{aligned} \tag{3.3}$$

The smallest value of x_3 for which the derivation of (3.3) is applicable is $x_3 = x_2$; for this choice of x_3, it follows from the first part of (3.3) that $H_2 = G_2 > H_0$. Hence when H_0 exceeds G_2, x_3 will be larger than x_2 and (3.3) is valid.

A similar argument applied to the trajectory 456 yields

$$\begin{aligned} 0 < 2(H_6 - H_4) &< -2F_1\{[2H_5]^{1/2} - [2(H_5 - G_2)]^{1/2}\} \\ &< \frac{-4F_1G_2}{\{[2H_5]^{1/2} + [2(H_5 - G_2)]^{1/2}\}} < \frac{-4F_1G_2}{\{(2H_6)^{1/2} + F_1\}}\,, \end{aligned} \tag{3.4}$$

where the last inequality in (3.4) is based on the assumptions $H_5 - G_2 > 0$ and $[2H_6]^{1/2} + F_1 > 0$. However, when $H_6 \geqq H_0$, these are both satisfied if H_0 is chosen so large that

$$(2H_0)^{1/2} \geqq (2G_2)^{1/2} - F_1\,. \tag{3.5}$$

For,

$$(2H_5)^{1/2} = (2H_6)^{1/2} - I_{56}(0) > (2H_6)^{1/2} + F_1 \geqq (2H_0)^{1/2} + F_1 \geqq (2G_2)^{1/2} > 0.$$

Therefore, assuming $H_6 \geqq H_0$ and (3.5), one has from Eqs. (3.3) and (3.4)

$$\begin{aligned} H_2 - H_4 < (H_2 - H_0) + (H_6 - H_4) &< -\frac{2F_1G_2}{(2H_0)^{1/2}} - \frac{2F_1G_2}{(2H_0)^{1/2} + F_1} \\ &= -2F_1G_2\left\{\frac{1}{(2H_0)^{1/2}} + \frac{1}{[(2H_0)^{1/2} + F_1]}\right\}. \end{aligned} \tag{3.6}$$

[1]The following analysis, which leads to inequality (3.3), sharper than the author's original result, is due to the reviewer.

4. Behaviour of the solution for $x \geqq x_2$. Since $I_{32}(G) = I_{43}(G) = F_3$ and $(H - G)_{32} > (H - G)_{43}$,

$$H_2 - H_3 = \int_{32} dH > \int_{43} dH = H_3 - H_4 ,$$
$$H_2 - H_4 = (H_2 - H_3) + (H_3 - H_4) < 2(H_2 - H_3). \tag{4.1}$$

Also,

$$I_{32}(G_2) = [2(H_2 - G_2)]^{1/2} - [2(H_3 - G_2)]^{1/2} < I_{32}(G)$$
$$= F_3 < I_{32}(G_3) = [2(H_2 - G_3)]^{1/2} < [2(H_2 - H_4)]^{1/2}. \tag{4.2}$$

From the first part of (4.2), it follows that

$$2(H_2 - G_2) < [2(H_3 - G_2) + F_3]^2,$$
$$2(H_2 - H_3) < F_3^2 + 2F_3[2(H_3 - G_2)]^{1/2} = F_3^2 + 2F_3[2(G_3 - G_2)]^{1/2}. \tag{4.3}$$

Combining with (4.1), we obtain

$$H_2 - H_4 < F_3^2 + 2F_3[2(G_3 - G_2)]^{1/2}. \tag{4.4}$$

Using the second part of (4.2), we have

$$\tfrac{1}{2}F_3^2 < H_2 - H_4 < F_3^2 + 2F_3[2(G_3 - G_2)]^{1/2}. \tag{4.5}$$

Using (3.6) and (4.5), we can now state:

Theorem 1. *When conditions* (A)-(E) *are satisfied and* H_0 *is so large that*

$$(2H_0)^{1/2} \geqq (2G_2)^{1/2} - F_1 , \tag{4.6}$$

$$F_3^2 \geqq -4F_1G_2\left\{\frac{1}{(2H_0)^{1/2}} + \frac{1}{[(2H_0)^{1/2} + F_1]}\right\}, \tag{4.7}$$

then $x_3 > x_2$ *and* $H_6 < H_0$.

As H_0 increases, H_2 increases. If $H_2 - H_4 \to \infty$, then, from (4.5), $F(\infty)\cdot G(\infty) = \infty$. On the other hand, $F(\infty) = \infty$ implies $H_2 - H_4 \to \infty$, by (4.5). If $G(\infty) = \infty$, then $H_2 (> G_3)$ approaches infinity as $x_3 \to \infty$. Hence by (3.3), $H_0 \to \infty$. Therefore condition (F) is equivalent to:

$$H_0 \to \infty \qquad \text{as} \qquad x_3 \to \infty .$$

Now suppose that the differential equation (1.2) is defined for $x < 0$ and that conditions corresponding to (A)-(F) are satisfied for $x \leqq 0$; specifically, assume that:

(A′) $f(x)$, $g(x)$ are continuous for $x \leqq 0$,
(B′) $f(x) < 0$ for $x_1' < x < 0$,
(C′) $f(x) > 0$ for $x < x_1'$,
(D′) $-F(-\infty) > 0$,
(E′) $-g(x) > 0$ for $x < 0$,
(F′) $-F(-\infty)G(-\infty) = \infty$.

If one replaces x by $-\xi$, eq. (1.2) becomes

$$\ddot{\xi} + f_1(\xi)\dot{\xi} + g_1(\xi) = 0,$$

where

$$f_1(\xi) = f(-\xi),$$

$$g_1(\xi) = -g(-\xi),$$

Then

$$f_1(\xi),\ g_1(\xi),\ F_1(\xi) = \int_0^{\xi} f_1(\xi)\,d\xi, \qquad G_1(\xi) = \int_0^{\xi} g_1(\xi)\,d\xi$$

satisfy conditions (A)-(F), although with x_1 replaced by $-x_1'$.

Now consider in the region $x \geqq 0$ a trajectory starting with initial energy H_0, returning to $x = 0$ with energy H_6, continuing into the region $x < 0$ (with initial energy H_6), and returning to $x = 0$ with energy H_{12}. Because of assumption (F), H_6 and H_{12} are monotone increasing functions of H_0 in the interval $0 \leqq H_0 < H_{12}$. According to Theorem 1, $H_6 < H_0$ for sufficiently large H_0. If H_6 were unbounded as $H_0 \to \infty$, then, applying Theorem 1 to the region $\xi = -x \geqq 0$, H_6 may be chosen so large that $H_{12} < H_6$. Hence for large H_0, $H_{12} < H_0$.

Now consider motions of small amplitude. When $x_3 \leqq x_2$, $H_0 < H_6$. A similar result holds for negative x. Therefore for small H_0, $H_{12} > H_0$.

Therefore H_{12} is a continuous, monotone-increasing function of H_0 for $0 \leqq H_0 < \infty$ such that $H_{12} - H_0$ changes sign. For some H_0, $H_{12} = H_0$, showing that there is a periodic solution of eq. (1.2). It is shown in Ref. 3 that $H_6 - H_0$ is monotone decreasing. Then $H_{12} - H_6$ and $H_6 - H_0$ are also monotone-decreasing and the periodic solution is unique.

Theorem 2. *If Eq.* (1.2) *satisfies conditions* (A)-(F) *and* (A′)-(F′), *there is a unique periodic solution.*

5. A counter-example. Let us denote a trajectory *recurrent* when $H_6 = H_0$. It is shown in Ref. 3 that if $f(x)$ is even and $g(x)$ odd relative to $x = 0$, a periodic characteristic is recurrent. It will now be shown that when condition (F) is not satisfied, there is, in general, no recurrent trajectory and therefore no periodic solution.

Consider the family of equations (1.2) defined by

$$F_\epsilon(x) = F(x), \qquad 0 \leqq x \leqq x_2,$$

$$= \epsilon F(x), \qquad x \geqq x_2, \qquad 0 < \epsilon \leqq 1$$

For every ϵ,

$$G_2 \leqq H_4 \leqq G(\infty) < \infty. \tag{5.1}$$

Now for every choice of H_4 satisfying (5.1), consider the trajectory 456 (Fig. 1) and that trajectory 012 for which $H_0 = H_6$ of the afore-mentioned part 456. Then $H_2 - H_4$ is a positive-valued, continuous function of H_4 on the closed interval (5.1); $H_2 - H_4$ therefore attains a minimum $\Delta > 0$ for some H_4.

If there were a recurrent solution for (1.2) for every ϵ, $0 < \epsilon \leqq 1$, then from (4.4),

$$0 < \Delta \leqq H_2 - H_4 < [F_\epsilon(x_3)]^2 + 2F_\epsilon(x_3)\{2[G(x_3) - G_2]\}^{1/2}$$

$$< \epsilon^2[F(\infty)]^2 + 2\epsilon F(\infty)\{2[G(\infty) - G_2]\}^{1/2}.$$

By choosing ϵ small enough, the inequality may be violated. Hence for such an ϵ, $H_0 \neq H_6$.

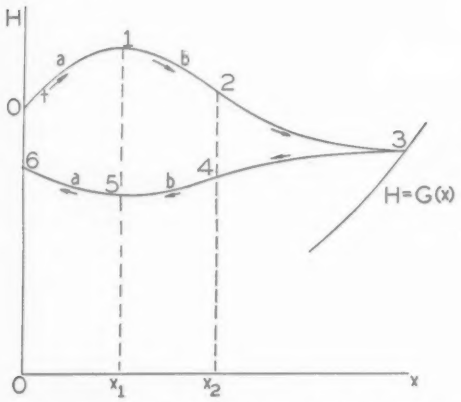

FIG. 1. Qualitative characteristics of the trajectory.

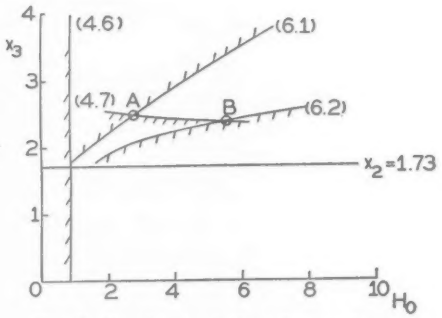

FIG. 2. Graphical solution of upper bounds.

6. Numerical bounds. It is useful, for application to specific problems, to develop numerical bounds to the amplitude x_3 of recurrent motion. Since $x_3 > x_2$, a lower bound is x_2. The following set of inequalities is useful, in some cases, for obtaining an upper bound to x_3 and H_0.

For $x_3 > x_2$, one has, using (3.3) and (4.3)

$$G_3 < H_0 + (H_2 - H_0) < H_0 - \frac{2F_1G_2}{(2H_0)^{1/2}}, \tag{6.1}$$

$$H_0 < (H_2 - H_3) + H_3 < \frac{F_3^2}{2} + F_3[2(G_3 - G_2)]^{1/2} + G_3. \tag{6.2}$$

A trajectory satisfying inequalities (4.6), (4.7), (6.1), and (6.2) has the property $H_0 > H_6$ and its maximum amplitude x_3 and initial ordinate H_0 furnish upper bounds to the corresponding characteristics of the periodic motion.

Example. Take $f(x) = -\frac{1}{2}(1 - x^2)$, $g(x) = x$.
Then,

$$F(x) = \frac{1}{2}\left(-x + \frac{x^3}{3}\right), \qquad G(x) = \frac{x^2}{2}, \qquad x_1 = 1, \qquad x_2 = 3^{1/2},$$

$$F_1 = -\frac{1}{3}, \qquad G_2 = \frac{1}{2}, \qquad G_2 = \frac{3}{2}.$$

The graphical solution of the inequalities is shown in Fig. 2. The curve of x_3 vs. H_0

lies between the curves labelled (6.1) and (6.2); the intersection with curve (4.7) defines a point whose coordinates are suitable as upper bounds. Referring to Fig. 2, it is seen that x_3 at point A is an upper bound to the amplitude and H_0 at point B is an upper bound to the initial energy. Thus

$$1.73 < x_3 < 2.50$$

and the average 2.11 of the bounds is in error by less than 23%.

List of References

1. Liénard, A.: *Étude des oscillations entretenues*, Revue Gén. de l'Electricité, t. XXIII (1928), pp. 901-946.

2. Levinson, N. and Smith, O. K.: *A general equation for relaxation oscillations*, Duke Math. Journal, 9, 382-403 (1942).

3. Lefschetz, S.: *Lectures on differential equations*, Princeton University Press, Princeton, N. J.

QUENCHING STRESSES IN TRANSPARENT ISOTROPIC MEDIA AND THE PHOTOELASTIC METHOD*

By R. C. O'ROURKE and A. W. SAENZ (*University of Michigan*)

Introduction. It is well known that when a transparent non-crystalline solid, such as glass, is heated to a uniform temperature T_1 and then rapidly quenched in a bath at temperature $T_0 (T_1 > T_0)$ there results a non-uniform stress distribution. Depending on T_1, one can divide these stresses into two distinct classes. If T_1 is sufficiently lower than the softening temperature of the material, then the stresses are transient, but if T_1 is sufficiently high (say 550° C for lime glass) then the quenching stresses are not transient but, on the contrary, remain permanently set into the glass. The latter are referred to in the literature as quenching or residual stresses. As is well known, the existence of such residual stresses is made manifest when the object is examined in polarized light, and by violent explosive characteristics of quenched objects when cut.† Each distribution of such stresses is characterized by a definite double refraction pattern. In 1841, F. E. Neumann[1] developed a general mathematical theory of the double refraction of light in non-uniformly heated isotropic solids. In turn the problem was studied theoretically by such men as Maxwell,[2] and Lord Rayleigh.[3]

The purpose of this paper is three-fold:

(1) To develop a mathematical theory of residual stresses based on a simple model

*Received July 11, 1949.

†For details on the background of the subject see E. G. Coker and L. N. G. Filon, *Treatise on photoelasticity*, Cambridge, 1931, §§ 332 and 333.

[1]F. E. Neumann, *Die Gesetze der Doppelbrechnung des Lichtes in comprimierten oder ungleichförmig unkrystallinischen Körpern*, Abh. Königl. Acad. Wiss. Berlin, Part II, pp. 1-254 (1841).

[2]J. C. Maxwell, *On the equilibrium of elastic solids*, Roy. Soc. Edin. Trans. **20**, 87-120 (1853).

[3]Lord Rayleigh, *On the stresses in solid bodies due to unequal heating and on the double refraction resulting therefrom*, VI-1, pp. 169-178, 1901; see also Arch. Neerl. (II) **5**, 32-42, (1900) and *Collected Papers*, vol. 4, p. 502.

of the quenching process. This theory is essentially a careful reinterpretation of the classical theory of thermal stresses, and is implicit in the work of Lord Rayleigh.[3]

(2) To justify, at least in a preliminary way, the use of the photoelastic method in the study of quenching stresses with particular emphasis on glass.

(3) To recognize the problem of determining the distribution of stresses photoelastically in symmetrically quenched spheres and cylinders as a simple problem in integral equations.

1. **Photoelastic preliminaries.** In this section the necessary photoelastic formulas used in the study of residual stresses in spheres and cylinders will be developed. The Maxwell-Neumann stress-optic law[4] states that the secondary principal stresses in the plane of the wave front are related to the retardation dR, suffered in traveling a distance dy, as follows:

$$dR = C(P - Q)\, dy \tag{1.01}$$

provided the direction of the principal stresses in the plane of the wavefront does not vary along the light path (Fig. 1). It is desired to calculate the integrated retardation suffered by a plane monochromatic wave in passing through symmetrically quenched spheres and cylinders. The wavefronts are parallel to the x-z plane (Fig. 1). For the integrated retardation one can write:

$$R(\eta) = C \int_{-\eta}^{+\eta} (P - Q)\, dy \tag{1.02}$$

(C is the stress-optic coefficient in Brewsters; y is measured in mm; $(P - Q)$ is the difference of the two principal stresses in the plane of the wavefront).

In the case of a long cylinder, whose surface is kept at T_0, the stress tensor in cylindrical coordinates assumes the following diagonal form:

$$|| \sigma_{ij} || = \begin{Vmatrix} \sigma_\rho(\rho) & 0 & 0 \\ 0 & \sigma_\phi(\rho) & 0 \\ 0 & 0 & \sigma_z(\rho) \end{Vmatrix}, \tag{1.03}$$

where σ_ρ, σ_ϕ, and σ_z are functions of ρ only. In order to obtain the principal stresses in the plane of the wavefront, one must transform the stress tensor (1.03) to appropriate cartesian axes (See Fig. 1) by means of a rotation about the z axis through an angle φ, given by

$$R_z(\phi) = \begin{Vmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{Vmatrix} \tag{1.04}$$

which transforms the stress tensor to:

[4]For the basic photoelastic concepts, see Coker and Filon, *loc. cit.*, Chapter III, especially paragraph 3 and 10.

$$\begin{Vmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{Vmatrix} = \begin{Vmatrix} \cos^2\phi\sigma_\rho + \sin^2\phi\sigma_\phi & \sin\phi\cos\phi(\sigma_\rho - \sigma_\phi) & 0 \\ \sin\phi\cos\phi(\sigma_\rho - \sigma_\phi) & \sin^2\phi\sigma_\rho + \cos^2\phi\sigma_\phi & 0 \\ 0 & 0 & \sigma_z \end{Vmatrix} \tag{1.05}$$

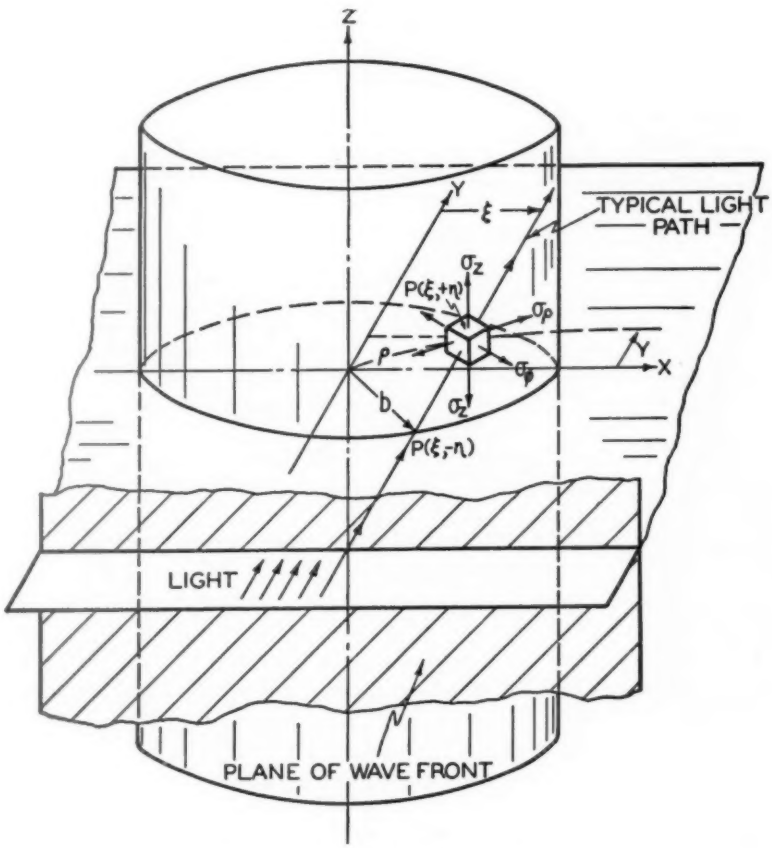

FIG. 1. Normal incidence on cylinder.

From (1.05) one readily identifies the principal stresses in the plane $x - z$ of the wave-front as

$$\begin{aligned} P &= \sigma_z , \\ Q &= \cos^2\phi\sigma_\rho + \sin^2\phi\sigma_\phi . \end{aligned} \tag{1.06}$$

Reversing roles of P and Q merely changes the sign of the retardation. This is equivalent to inverting the Babinet Compensator.[5] Thus the total retardation becomes:

$$R(\eta) = C \int_{-\eta}^{+\eta} \{\sigma_z - \cos^2 \phi\sigma_\rho - \sin^2 \phi\sigma_\phi\}\, dy. \tag{1.07}$$

Now, this integral can be simplified because of the inherent symmetry involved in the study of spheres and cylinders. One expresses $\cos^2 \varphi$, $\sin^2 \varphi$, and y in terms of ξ (Fig. 1)

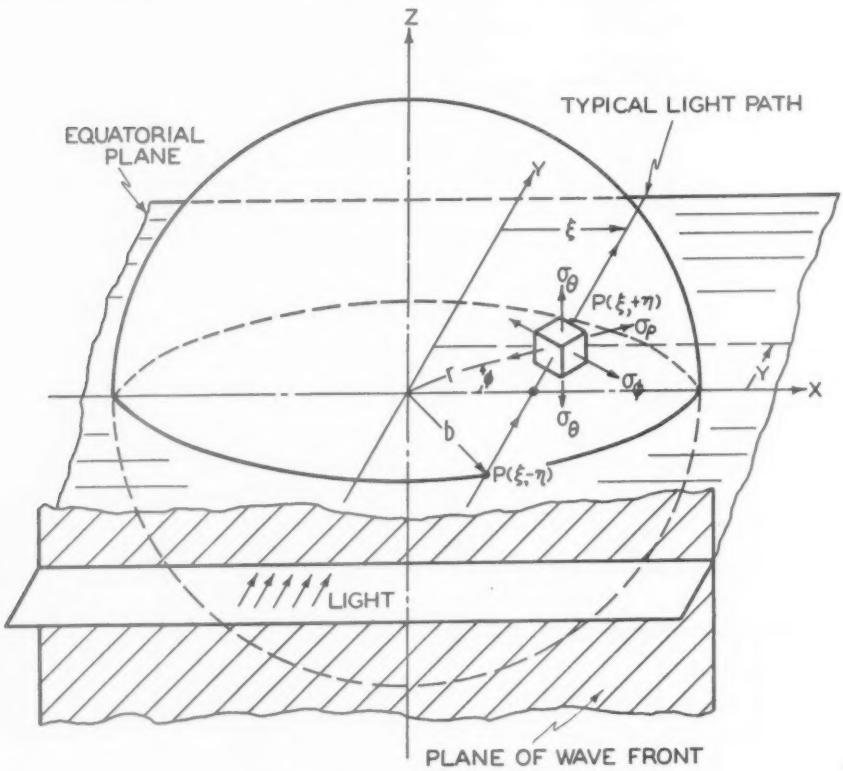

FIG. 2. Normal incidence on sphere.

and ρ, and notes the fact that σ_ρ , σ_ϕ and σ_z are functions of ρ only. Then equation (1.07) becomes:

$$R(\xi) = 2C \int_{\xi}^{b} \left\{(\sigma_z - \sigma_\phi) + (\sigma_\phi - \sigma_\rho)\frac{\xi^2}{\rho^2}\right\} \frac{\rho\, d\rho}{(\rho^2 - \xi^2)^{1/2}}\,. \tag{1.08}$$

This equation is of fundamental importance for the study of residual stresses in cylinders and spheres. It can be studied in two different ways, both of which are treated in detail below. The function $R^*(\xi)$, (an asterisk will signify that the function is obtained

[5]Coker and Filon, *loc. cit.* paragraphs 1 and 36.

experimentally), is easily obtained with a suitable Babinet compensator. (1.08) reduces to an Abel integral equation with the aid of appropriate substitutions, thereby enabling one to give an exact solution for each of the three stresses, σ_ρ , σ_ϕ , σ_z . (See Section III). The second way of treating (1.08) is to employ a model of the quenching process (Sect. II), and calculate σ_ρ , σ_ϕ , and σ_z , thereby finding a theoretical retardation pattern $R(\xi)$ which can then be compared with experiment. The case of the sphere lends itself to an analogous treatment. One measures $R^*(\xi)$ along the equatorial plane (See Fig. 2) where no rotation in the plane of the wavefront occurs. The stress tensor is diagonal in spherical coordinates r, ϑ, φ, for a sphere quenched in the same way as the cylinder above. It assumes the form:

$$|| \sigma_{ij} || = \begin{Vmatrix} \sigma_r(r) & 0 & 0 \\ 0 & \sigma_\vartheta(r) & 0 \\ 0 & 0 & \sigma_\varphi(r) \end{Vmatrix}; \qquad \sigma_\vartheta = \sigma_\phi \,. \tag{1.09}$$

The retardation integral for the sphere becomes:

$$R(\xi) = 2C\xi^2 \int_\xi^b (\sigma_\phi - \sigma_r) \frac{dr}{r(r^2 - \xi^2)^{1/2}} \,. \tag{1.10}$$

2. Applicability of photoelastic method in the study of quenching stresses. In Sec. 3 a method will be developed which reduces the calculation of residual stresses in symmetrically quenched spheres and cylinders to the inversion of Abel's integral equation. For the cylinder, this method will depend upon the assumption that residual stresses are elastic stresses; this assumption is implied in the "sum rule." The "sum rule" is the relation $\sigma_z(\rho) = \sigma_\rho(\rho) + \sigma_\phi(\rho)$ which is characteristic of thermoelastic stresses in symmetrically quenched circular cylinders (hollow and solid). The necessity for making this assumption lies in the fact that the method being proposed uses *relative* retardations $R^*(\xi)$ (Babinet Compensator) instead of *absolute* retardations (Mach-Zehnder interferometer). If *absolute* retardations are measured, then one can invert the appropriate Abel integral equations of Sec. 3 without the "sum rule" ansatz. (This will be reported on elsewhere.) However, for practical purposes, it is much more expedient to measure *relative* retardations since, as will be shown in Sec. 3, one obtains the complete distribution of residual stresses in symmetrically quenched cylinders from a single photograph of $R^*(\xi)$.

It is therefore necessary to give some justification for the "sum rule" ansatz in the case of residual stresses of the type being considered. There are at least four exact methods for showing the elastic nature of these residual stresses, two of which will be explained now. First of all, Eq. (1.08) can be written in the following form:

$$R^*(\xi) = 2C \int_\xi^b \sigma_z \frac{\rho \, d\rho}{(\rho^2 - \xi^2)^{1/2}} \,, \tag{2.01}$$

since equilibrium demands that

$$\int_0^\eta Q \, dy = 0.$$

Now if one uses the "sum rule" and the equilibrium equation

$$\frac{d}{d\rho}(\rho\sigma_\rho) - \sigma_\phi = 0$$

Eq. (2.01) becomes

$$R^*(\xi) = 2C \int_\xi^b \frac{d}{d\rho}(\rho^2\sigma_\rho) \frac{d\rho}{(\rho^2 - \xi^2)^{1/2}} \cdot \tag{2.02}$$

From (2.02), using Dirichlet's principle for inverting orders of integration and the fact that $\sigma_\rho(a) = \sigma_\rho(b) = 0$, one easily derives the interesting relation

$$\int_a^b R^*(\xi)\, d\xi = 0 \tag{2.03}$$

For those residual stress patterns which have been studied, (2.03) is very accurately satisfied. Secondly, it is of further interest to see if one can construct a model of the quenching stresses from which a theoretical expression for $R(\xi)$ can be computed and compared with the experimental pattern $R^*(\xi)$. This is indeed possible and will be briefly discussed here. First of all, one assumes the existence of a "state of ease" at a temperature T^*, such that for $T > T^*$ all stresses in the body are negligibly small compared to the stresses for $T \ll T^*$ (of order $10^{kg/}$ mm^2 for lime glass). Secondly, one neglects the temperature variation of the elastic and thermal constants. The striking agreement obtained so far between theory and experiment seems to indicate that the model is essentially correct. Many experimental investigations have revealed the fact that the properties of amorphous materials, such as glass, change rather abruptly at a certain critical temperature.[5] The proposed model merely idealizes this fact by making the elastic constants change discontinuously at the critical temperature T^*.

These assumptions readily lead to the following rule:

> To calculate the residual stresses in a body (A) quenched from the uniform temperature T_1 to the temperature T_0 of a bath ($T_1 > T_0$), one calculates the thermal stresses at time $t = t^*$ ($T = T^*$) in an elastic body (B), with the same elastic and thermal constants as (A), whose initial temperature is T_1 and whose surface beginning at $t = 0$ is maintained at T_1 . The stresses in (B) at $t = t^*$ are the negative of the residual stresses at $t \to \infty$ in (A). (2.04)

Since the problems of thermoelastic stresses in spheres and cylinders have been dealt with thoroughly in the literature, it is not necessary to discuss them here. Essentially, one takes the pertinent solutions (say from Timoshenko[6]), replaces t by t^*, changes the signs of the stresses, according to (2.04), and substitutes into (1.08) and (1.10) to obtain the theoretical retardation patterns. The theoretical retardation patterns for cylinders and spheres have been calculated in this fashion and are given below:

---

[5] R. Houwink, *Elasticity, plasticity, and structure of matter*, Cambridge, 1940, Chap. 3 and 6.

[6] S. Timoshenko, *Theory of elasticity*, New York, 1934, Chap. 11.

Solid Cylinder (radius b):

$$R(\lambda) = -\frac{4C\alpha E(T_1 - T_0)}{1-\nu} \cdot b \cdot \sum_{n=1}^{\infty} \frac{\exp(-\chi\xi_n^2 t^*)}{\xi_n^2 b^2 J_1(\xi_n b)} R(n;\lambda)$$

$$R(n;\lambda) = -J_1(\xi_n b)(1-\lambda^2)^{1/2} + \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{(-)^{k+l}(2l+1)\cdot(\xi_n b)^{2k+2l}}{4^{k+l}\cdot k!(k+2l+1)!} \tag{2.05}$$

where

$$\cdot(1-\lambda^2)^{k+l-1/2} \cdot \frac{J_{2l+1}(\lambda\xi_n b)}{\lambda}\left\{\frac{1-\lambda^2}{2(k+l)+1} + \frac{2(k+l)\lambda^2}{2(k+l)-1}\right\}$$

ξ_n are positive roots of $J_0(\xi b) = 0$ and $\lambda = \xi/b$ (Fig. 1)

In practice t^* is sufficiently large so that only the "shape factor" $R(1, \lambda)$ need be calculated (see Fig. 3):

(B) Hollow Cylinder: $(a \leq \rho \leq b)$

$$R(\lambda) = -2C\frac{\alpha E}{1-\nu}\sum_{n=1}^{\infty}\exp(-\chi\xi_n^2 t^*)\frac{A_n}{\xi_n}\left\{\frac{2a^2 Z_1(\xi_n b) - 2abZ_1(\xi_n a)}{b^2-a^2}\right.$$

$$\cdot(1-\lambda^2)^{1/2} + \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-)^{k+l+1}(2l+1)\cdot(\xi_n b)^{2k+2l}}{4^{k+l}k!(2l+k+1)!}\frac{Z_{2l+1}(\lambda\xi_n b)}{\lambda} \tag{2.06}$$

$$\left.\cdot(1-\lambda^2)^{k+l-1/2}\left\{\frac{1-\lambda^2}{2(k+l)+1} + \frac{2(k+l)\lambda^2}{2(k+l)-1}\right\}\right,$$

where

$$A_n = \frac{\pi^2}{2}(T_1 - T_0)\xi_n \frac{J_0^2(\xi_n a)}{J_0^2(\xi_n b) - J_0^2(\xi_n a)}\{bZ_1(\xi_n b) - aZ_1(\xi_n a)\}$$

$$Z_0(x) = J_0(x)N_0(\xi_n b) - N_0(x)J_0(\xi_n b)$$

ξ_n are roots of $Z_0(\xi a) = 0$

(C) Sphere (radius b)

$$R(\lambda) = -\frac{4C\alpha E}{1-\nu}(T_1 - T_0)\left(\frac{\pi^3}{2}\right)^{1/2} \cdot b \cdot \sum_{n=1}^{\infty}(-)^n n \exp\left(-\frac{n^2\pi^2\chi t^*}{b^2}\right)R(n;\lambda), \tag{2.07}$$

where

$$R(n;\lambda) = \lambda^2 \sum_{k=0}^{\infty}\frac{\pi^k(2n)^{k-1/2}k!}{(2k+1)!}(1-\lambda^2)^{k+1/2}J_{k+5/2}(n\pi)$$

The leading terms $R(1;\lambda)$ for (2.05) and (2.07) are plotted in figures 3 and 4 respectively. The agreement between these "first approximations" $R(1;\lambda)$ and the corresponding experimental patterns $R^*(\lambda)$ obtained so far furnishes strong argument for the applicability of the photoelastic method in the determination of the type of quenching stresses under study here.

3. Photoelastic determination of quenching stresses in spheres and cylinders. Equation (2.02) is Abel's integral equation so that the problem of finding the distribution of residual stresses in symmetrically quenched cylinders reduces to the inversion of this

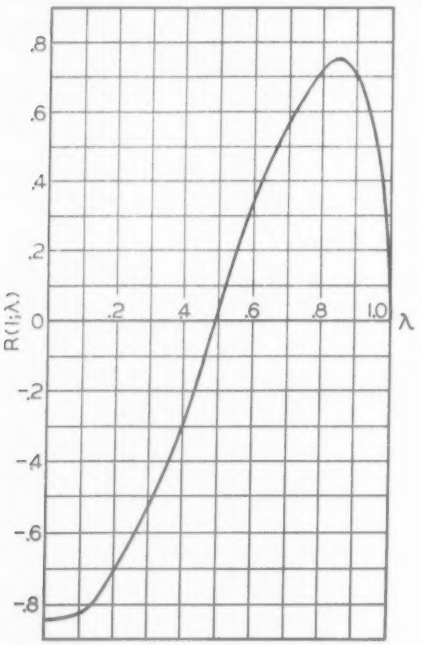

FIG. 3. Characteristic photoelastic curve for quenched solid cylinder.

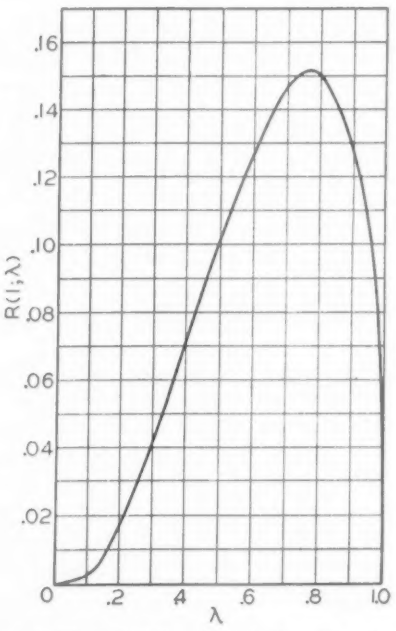

FIG. 4. Characteristic photoelastic curve for quenched glass sphere.

integral equation, which is well known. The complete solution for cylinders can be written in the form:

$$\sigma_\rho(\rho) = -\frac{1}{\pi C}\frac{1}{\rho^2}\int_\rho^b \frac{\xi R^*(\xi)\,d\xi}{(\xi^2-\rho^2)^{1/2}},$$

$$\sigma_\phi(\rho) = \frac{d}{d\rho}(\rho\sigma_\rho(\rho)), \tag{3.01}$$

$$\sigma_z(\rho) = \sigma_\rho(\rho) + \sigma_\phi(\rho).$$

For the case of spheres, one immediately recognizes (1.10) as Abel's integral equation without the necessity of a "sum rule" (which does not exist for the corresponding thermoelastic stresses). A similar inversion yields:

$$\sigma_r(r) = -\frac{2}{\pi C}\int_r^b \frac{R^*(\xi)\,d\xi}{\xi(\xi^2-r^2)^{1/2}},$$

$$\sigma_\vartheta(r) = \sigma_\phi(r) = \frac{r}{2}\frac{d}{dr}\sigma_r + \sigma_r\,. \tag{3.02}$$

Thus, the entire problem reduces to quadratures. $R^*(\xi)$ is an analytic function of the light-path variable η (or $(b^2 - \xi^2)^{1/2}$). Thus one can write

$$R^*(\xi) = \sum_{k=1}^{\infty} a_k(b^2 - \xi^2)^{k/2} \tag{3.03}$$

If (3.03) is substituted into (3.01) the quadratures can be performed and one finds for cylinders:

$$\rho^2\sigma_\rho(\rho) = -\frac{\Gamma(1/2)}{2\pi C}\sum_{k=1}^{\infty} a_k \frac{\Gamma(k+1)}{\Gamma(k+3/2)}(b^2 - \rho^2)^{(k+1)/2} \tag{3.04}$$

Correspondingly for spheres equation (2.07) suggests the following ansatz:

$$R^*(\xi) = \sum_{k=1}^{\infty} c_k(b^2 - \xi^2)^{k/2}\xi^2 \tag{3.05}$$

and (3.02) yields

$$\sigma_r(r) = -\frac{1}{\pi C}\Gamma(1/2)\sum_{k=1}^{\infty} c_k \frac{\Gamma(k+1)}{\Gamma(k+3/2)}(b^2 - r^2)^{(k+1)/2} \tag{3.06}$$

The coefficients a_κ , c_κ are easily obtained by fitting $R^*(\xi)$ by the method of least squares.

From the above results, it is seen that the method in this section rests on the "sum rule" for the case of concentric and solid cylinders, while in that of spheres no such assumption has to be made. It is thus clear that equations (3.01) and (3.02) are independent of the detailed nature of the boundary conditions of the problem, and that they are therefore applicable when large radiation losses are present which cannot be accounted for by the simple Newton law of cooling, a situation which would preclude the use of the method of Sec. 2, from a practical standpoint.

The results obtained so far employing the procedures of Sec. 3 are of technical importance in locating exposed tension regions (i.e., unstable regions) in quenched cylinders and spheres. In such work one desires to know the sign of the boundary stresses, their absolute magnitude being of secondary importance. One could also gain rough ideas concerning residual stresses in non-transparent polycrystalline materials, such as steel, by simply changing the elastic and thermal constants appropriately and using ideas of similarity.

STEFAN-LIKE PROBLEMS*[1]

BY G. W. EVANS II, E. ISAACSON, AND J. K. L. MACDONALD[2] (*New York University*)

Introduction. Consider the recrystallization of an infinite metal slab, as for example, the change from α to β crystals when iron is heated through the temperature 1643°F. Accompanying this change of crystalline form is a latent heat of recrystallization. If the slab is originally of uniform crystalline structure and is heated from the front face through the critical temperature, new crystals appear at the heated face. If the front face is planar and is heated uniformly and the back face is insulated, the interface between the two types of crystals is planar. For the determination of stresses occurring in the metal, it would be necessary to determine the history of the interface as it moves inward. To determine the history (a space-time curve) of this interface, it is necessary to solve the heat equation, in this case, a linear homogeneous parabolic equation, with appropriate boundary conditions and with a suitable discontinuity in the first order space derivatives across the unknown interface.

In each of the several problems to be discussed here, we will consider the recrystallization of a metal slab of uniform thickness, L, bounded by the planes $x = 0$ and $x = L$. We assume that a uniform heat source covers the front face, $x = 0$, and that a perfect insulator covers the back face, $x = L$, if L is not infinite. Thus, the problems will involve only one space dimension. Furthermore, the coefficient of thermal diffusivity, α^2, is assumed to be a constant for a given type of crystal, and recrystallization is assumed to take place at a temperature $u = 0$. Since in each problem the object is to find the propagation of the interface through the slab, the choice of the critical temperature as zero in no way affects the results.

To determine the type of discontinuity across the interface, assume that the recrystallization has already occurred to a distance x where $0 < x < L$. The subscript 1 after a symbol will be used to indicate that the symbol refers to the region of the original crystalline form and the subscript 2 to indicate that the symbol refers to the recrystallized region. The equation of heat balance across the interface as it moves a distance dx then is, after division by dt,

$$\rho_1 H \dot{x}(t) = (k_2 u_{2x} - k_1 u_{1x}) \mid_x = x(t), \tag{1}$$

where $x = x(t)$ is the space-time curve along which the interface moves, $\dot{x}(t) = dx/dt$, ρ is the density, H is the latent heat of recrystallization, and the k_i are the coefficients of thermal conductivity. The notation used is $\partial u_i/\partial x = u_{ix}$ for $i = 1, 2$. The temperature distribution in each region is given by the differential equation

$$u_{it} = \alpha_i^2 u_{ixx} \tag{2}$$

---

*Received Oct. 14, 1949.

[1]a) The term "Problem of Stefan" has recently been used for problems similar to the ones treated here by L. I. Rubinstein, A. B. Datsev, and others.

b) The authors express their indebtedness to the members of the Institute of Mathematics and Mechanics and especially to Professors R. Courant, B. Friedman, and J. Keller.

[2]Deceased. The practical results of this paper were obtained prior to the untimely death of our inspiring co-author J. K.L. MacDonald. The surviving authors shall endeavor to complete, in the mathematical sense, the material covered in this note.

1. The first problem to be considered will illustrate a method of solution for all. In this problem the slab has any thickness L, $0 < L \leq \infty$. The slab is considered to have been heated uniformly to the critical temperature, 0, and then a constant heat source is applied to the front face of the slab. This may be represented by the following mathematical problem:

Find $x = x(t)$ where the temperature distribution satisfies the equations[3]

$$u_t = \alpha^2 u_{xx} \qquad \text{for } 0 < x < x(t), \tag{3}$$

$$u = 0 \qquad \text{for all } x \geq x(t) \tag{4}$$

with the boundary conditions

$$A\,\dot{x}(t) = u_x[x(t), t] \tag{5}$$

$$x(0) = 0 \tag{6}$$

$$u_x(0, t) = g, \tag{7}$$

where g is a constant. The condition (5) is a consequence of equation (1) where $A = \rho H/k$.

We obtain a power series representation for $x = x(t)$ about $t = 0$ by assuming u expressible in a power series about $x = 0, t = 0$ for all u in the region $0 \leq x \leq x(t)$, i.e., any discontinuities in the derivatives of u are assumed to occur only on crossing the curve $x = x(t)$.

By differentiation of (4) and (5) along the curve $x = x(t)$ and of (3) in the region $0 \leq x \leq x(t)$ and use of (7), all derivatives of $x(t)$ may be determined at the origin. $x(t)$ may then be written as the following Taylor series:

$$x(t) = \frac{g}{A}\,t - \frac{1}{2!}\frac{g^3}{\alpha^2 A^3}\,t^2 + \frac{5}{3!}\frac{g^5}{\alpha^4 A^5}\,t^3 - \frac{51}{4!}\frac{g^7}{\alpha^6 A^7}\,t^4 + \frac{827}{5!}\frac{g^9}{\alpha^8 A^9}\,t^5 - \cdots \tag{8}$$

The coefficients of the series have not been calculated beyond those given. They may be found from the following relations:

Let $u = \sum_{i,j=0}^{\infty} a_{ij} x^i t^j$ and $x = \sum_{i=0}^{\infty} C_i t^i$, then,

$$\sum_{i,j=0}^{\infty} (i+1) a_{i+1,j} \Big(\sum_{n=0}^{\infty} C_n t^n \Big)^i t^j = A \sum_{n=0}^{\infty} (n+1) C_{n+1} t^n, \tag{9}$$

$$\sum_{i,j=0}^{\infty} a_{ij} \Big(\sum_{n=0}^{\infty} C_n t^n \Big)^i t^j = 0, \tag{10}$$

$$a_{i+2,j} = \frac{1}{\alpha^2} \frac{(j+1)}{(i+1)(i+2)}\, a_{i,j+1} \tag{11}$$

with $a_{1,0} = g$ and $C_0 = 0$.

[3]We have adopted the convention of using the letter u without subscripts when the temperature is identically 0 in the region 1 (i.e., to the right of the discontinuity).

In the event that the heat input, g, is no longer a constant but an analytic function of time, $g = g(t)$, we find in a similar way the interface curve to be:

$$x(t) = C_1 t - \frac{C_2}{2!} t^2 + \frac{C_3}{3!} t^3 - \frac{C_4}{4!} t^4 + \frac{C_5}{5!} t^5 - \cdots , \tag{12}$$

where

$$C_1 = \frac{g(0)}{A} ,$$

$$C_2 = \frac{g^3(0)}{\alpha^2 A^3} - \frac{\dot{g}(0)}{A} ,$$

$$C_3 = \frac{5g^5(0)}{\alpha^4 A^5} - \frac{6\, \dot{g}(0) g^2(0)}{\alpha^2 A^3} + \frac{\ddot{g}(0)}{A} ,$$

$$C_4 = \frac{51 g^7(0)}{\alpha^6 A^7} - \frac{75\, \dot{g}(0)\, g^4(0)}{\alpha^4 A^5} + \frac{10\, \ddot{g}(0) g^2(0)}{\alpha^2 A^3} + \frac{15\, [\dot{g}(0)]^2\, g(0)}{\alpha^2 A^3} - \frac{\dddot{g}(0)}{A} ,$$

$$C_5 = \frac{827\, g^9(0)}{\alpha^8 A^9} - \frac{1438\, \dot{g}(0) g^6(0)}{\alpha^6 A^7} + \frac{175\, \ddot{g}(0) g^4(0)}{\alpha^4 A^5} + \frac{525\, [\dot{g}(0)]^2 g^3(0)}{\alpha^4 A^5} - \frac{15\, \dddot{g}(0)\, g^2(0)}{\alpha^2 A^3} - \frac{15\, [\dot{g}(0)]^3}{\alpha^2 A^3} - \frac{60\, \dot{g}(0)\, \ddot{g}(0)\, g(0)}{\alpha^2 A^3} + \frac{\ddddot{g}(0)}{A} .$$

The series of Eq. (12) reduces to that of Eq. (8) for $g(t)$ a constant, $g(0) = g$.

2. To generalize the treatment of Sec. 1 to cover the case of non-analytic boundary and initial conditions, we employ Laplace transformations. Furthermore, we believe that the questions of uniqueness and existence will be answered by these methods. In this part, we derive the integral equation which is satisfied by the interface curve $x = x(t)$. To this end, the discontinuity curve is considered in the form $t = f(x)$, and for greater generality we consider the heat input, $g = g(t)$, to be a function of time. The problem is to solve the differential equation

$$u_t(x, t) = \alpha^2 u_{xx}(x, t) \qquad \text{for} \qquad t > f(x) \tag{13}$$

with the boundary conditions

$$u(x, t) = 0, \qquad \text{for } t \leq f(x) \tag{14}$$

$$u_x(0, t) = g(t), \tag{15}$$

$$f(0) = 0, \tag{16}$$

$$u_x[x, f(x)] = \frac{A}{f'(x)} , \tag{17}$$

where $f' = df/dx$.

The Laplace transform for $u(x, t)$ is defined as

$$u^{\square}(x, s) = \int_0^{\infty} u(x, t)\, e^{-st}\, dt = \int_{f(x)}^{\infty} u(x, t)\, e^{-st}\, dt.$$

Application of the transform to (13) furnishes the transformed differential equation

$$\left[\frac{d^2}{dx^2} - \frac{s}{\alpha^2}\right] u^{\square} = -A\, e^{-sf(x)} \tag{18}$$

with the boundary conditions:

$$u_x^{\square}(0, s) = g^{\square}(s), \qquad \operatorname*{Lim}_{x\to\infty}\{u^{\square}(x, s)\} = 0.$$

The solution of Eq. (18) is

$$u^{\square}(x, s) = \frac{A\alpha}{2(s)^{1/2}} \int_0^{\infty} \exp\{-sf(\xi)\}\left[\exp\left\{-\frac{s^{1/2}}{\alpha} \mid x - \xi \mid\right\} + \exp\left\{-\frac{s^{1/2}}{\alpha}(x + \xi)\right\}\right] d\xi$$

$$- \frac{\alpha g^{\square}(s)}{s^{1/2}} \exp\left\{-\frac{s^{1/2}}{\alpha} x\right\},$$

and the inverse transform of this equation is

$$u(x, t) = \frac{A\alpha}{2} \int_0^{\infty} \frac{1}{\{\pi[t - f(\xi)]\}^{1/2}}\left[\exp\left\{\frac{-(x - \xi)^2}{4\alpha^2[t - f(\xi)]}\right\}\right.$$

$$\left. + \exp\left\{\frac{-(x + \xi)^2}{4\alpha^2[t - f(\xi)]}\right\}\right] d\xi \qquad (19)$$

$$_{t>f(\xi)}$$

$$- \alpha \int_0^t \frac{g(t - \tau)}{(\pi\tau)^{1/2}} \exp\left\{\frac{-x^2}{4\alpha^2\tau}\right\} d\tau.$$

Once $t = f(x)$ is known, Eq. (19) gives the temperature distribution in that region which has been recrystallized. Now, applying the condition (14) in the form $u[x, f(x)] = 0$, Eq. (19) becomes

$$2 \int_0^{f(x)} \frac{g(t - \tau)}{\tau^{1/2}} \exp\left\{\frac{-x^2}{4\alpha^2\tau}\right\} d\tau = A \int_0^{\infty} \frac{1}{[f(x) - f(\xi)]^{1/2}}\left[\exp\left\{\frac{-(x - \xi)^2}{4\alpha^2[f(x) - f(\xi)]}\right\}\right.$$

$$\left. + \exp\left\{\frac{-(x + \xi)^2}{4\alpha^2[f(x) - f(\xi)]}\right\}\right] d\xi \qquad (20)$$

$$_{f(x)>f(\xi)}\ .$$

Equation (20) is an integral equation for the determination of $t = f(x)$. To solve Eq. (20), we return to the designation $x = x(t)$. To do this let $t = f(x)$, $\tau = f(\xi)$, $x = x(t)$, and $\xi = x(\tau)$. This change is permissible if $t = f(x)$ is monotonically increasing. After making these substitutions in the right hand side of Eq. (20), replace $t - \tau$ by τ' and replace τ' by τ in the right hand side. Equation (20) becomes

$$2 \int_0^t \frac{g(t - \tau)}{\tau^{1/2}} \exp\left\{\frac{-x^2(t)}{4\alpha^2\tau}\right\} d\tau = A \int_0^t \frac{\dot{x}(t - \tau)}{\tau^{1/2}}\left[\exp\left\{\frac{-[x(t) - x(t - \tau)]^2}{4\alpha^2\tau}\right\}\right.$$

$$\left. + \exp\left\{\frac{-[x(t) + x(t - \tau)]^2}{4\alpha^2\tau}\right\}\right] d\tau \qquad (21)$$

$$_{\tau>0}\ .$$

A power series solution for $x = x(t)$ of equation (21) may be determined as follows: Let

$$x(t) = C_1 t + C_2 t^2 + \cdots , \tag{22}$$

and assume $g(t)$ a constant, g. The coefficients C_1 and C_2 are determined by using (22) to expand the integrands of equation (21) in powers of t (i.e., $x(t - \tau)$ is expanded about $\tau = 0$). We find that in the limit, as t tends to zero, equation (20) becomes

$$C_1 = \frac{g}{A} .$$

Now, replacing $x(t)$ by $gt/A + C_2 t^2$ and $x(t - \tau)$ by

$$x(t - \tau) = x(t) - \tau \dot{x}(t) + \tau^2 \frac{\ddot{x}(t)}{2!}$$

in equation (21) and expanding about $t = 0$, Eq. (21) becomes

$$4gt^{1/2} - \frac{2g^2\pi^{1/2}}{\alpha A} t + \frac{g^3}{A^2\alpha^2} = 4gt^{1/2}$$

$$+ \frac{16\,A}{3} C_2 t^{3/2} - \frac{2g^2\pi^{1/2}}{\alpha A} t + \frac{11}{3} \frac{g^3}{A^2\alpha^2} t^{3/2}$$

which yields

$$C_2 = -\frac{1}{2} \frac{g^3}{\alpha^2 A^3} .$$

It is noticed that the coefficients determined in this manner are the same as the first two of equation (8). Similarly, when $g(t)$ is an analytic function of t, the first two coefficients, C_1 and C_2 , of equation (22) have been determined and agree with those of equation (12). The determination of the remaining coefficients by this method becomes extremely tedious.

3. Another problem which involves the determination of the curve of recrystallization would be to prescribe an initial temperature distribution in a metal slab of infinite thickness ($L = \infty$). The problem is then to determine the curve $x = x(t)$ where the temperature distribution is given by

$$u_{2_t}(x, t) = \alpha_2^2 u_{2_{xx}}(x, t) \qquad \text{for } 0 < x < x(t), \tag{23}$$

$$u[x(t), t] = 0, \tag{24}$$

$$u_{1_t}(x, t) = \alpha_1^2 u_{1_{xx}}(x, t) \qquad \text{for } x(t) < x \leq \infty \tag{25}$$

with the boundary conditions

$$k_2 u_{2_x}[x(t), t] - k_1 u_{1_x}[x(t), t] = \rho H \dot{x}(t) = B \dot{x}(t), \tag{26}$$

$$u_1(x, 0) = \phi(x) \qquad \text{where} \qquad \phi(x) \leq 0 \qquad \text{and} \qquad \phi(0) = 0, \tag{27}$$

$$u_{2_x}(0, t) = g(t). \tag{28}$$

In this problem $g(t)$ is assumed to be the same as in the previous problem and $\phi(x)$ is assumed to be expressible in a Taylor expansion about $x = 0$ with an infinite radius of convergence. The method of solution is again the same as that of Part I and the series representing $x(t)$ is given below through the second power of t.

$$\begin{aligned} x(t) = {} & \frac{1}{B}\,[k_2 g(0) - k_1\phi'(0)]t \\ & + \frac{1}{2!B}\left\{k_2\left[-\frac{g(0)}{\alpha_2^2 B^2}\,(k_2 g(0) - k_1\phi'(0))^2\right.\right. \\ & \left. + g'(0)\right] - k_1\left[\frac{\phi''(0)}{B}\,(k_2 g(0) - k_1\phi'(0))\right. \\ & \left.\left. + \alpha_1^2\,\phi'''(0)\right]\right\} t^2 + \cdots . \end{aligned} \tag{29}$$

This problem is reducible to the second problem of Part I by taking $\phi(x) = 0$, and the coefficients in Eq. (29) then agree with those of Eq. (12).

Comments. The boundary condition (4), $u = 0$ for $x \geq x(t)$, used in the first two problems is not as specialized as may first appear. In most practical cases the metal is initially at a uniform temperature, say room temperature. If a metal slab, initially at a uniform temperature, is heated at one face and is insulated at the other, the temperature gradient remains zero for sufficiently small L. That is, the temperature rises uniformly until the critical temperature is reached providing the slab is thin. This result is seen if one lets $u = U > 0$ be the critical temperature and considers the slab initially to have the temperature distribution $u(x, 0) = 0$. One can then show that the temperature distribution is independent of the space coordinate as long as $u(x, t) < U$. The temperature distribution in the slab of thickness L is given by the differential equation

$$u_t = \alpha^2 u_{xx} \tag{30}$$

with the boundary conditions

$$u_x(0, t) = g(t), \tag{31}$$

$$u_x(L, t) = 0, \tag{32}$$

$$u(x, 0) = 0. \tag{33}$$

The Laplace transform of (30) is the equation

$$u_{xx}^{\square}(x, s) = \frac{s}{\alpha^2}\,u^{\square}(x, s),$$

the solution of which is

$$u^{\square}(x, s) = -\frac{\alpha g^{\square}(s)}{s^{1/2}}\,\frac{\cosh\,(L - x)s^{1/2}/\alpha}{\sinh\,Ls^{1/2}/\alpha} \tag{34}$$

for the transposed boundary conditions (31) and (32). Since the slab is considered thin,

replace $\cosh (L - x)\, s^{1/2}/\alpha$ and $\sinh Ls^{1/2}/\alpha$ by the first terms of their expansions giving

$$u^{\square}(x, s) = -\frac{\alpha^2}{sL}\, g^{\square}(s). \tag{35}$$

In (35), $u^{\square}(x, s)$ is the transform of $u(x, t)$, $g^{\square}(s)$ of $g(t)$ and $-\alpha^2/(sL)$ is the transform of $-\alpha^2/L$; therefore, $u(x, t)$ may be written as the following convolution

$$u(x, t) = -\int_0^t \frac{\alpha^2}{L}\, g(y)\, dy.$$

This expression shows $u(x, t)$ to be a function only of the variable t, and is therefore a constant with respect to x.

Discussion of the curves of equations (8), (12), and (29). a) Figure 1 shows the degree of agreement of the first five sums of the series of equation (8) where $x_1(t)$ represents

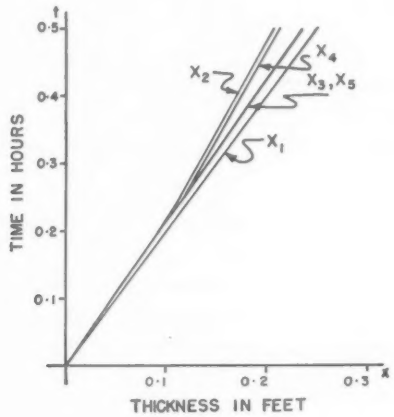

FIG. 1.

the first term on the right-hand side of (8), $x_2(t)$ represents the first plus the second, $x_3(t)$ represents the first three terms, etc. The calculations for this figure are based on the following data: $k = 25$ BTU/(hr.)(sq. ft.)(°F/ft.), $c_p = 14$ BTU/(lb.)(°F), $\rho = 480$ lbs./cu. ft., $H = -6.76$ BTU/lb., and $g = -65$ BTU/(sq. ft.)(hr.). This data is approximately that for a cast iron.

b) Figure 2 shows a comparison of the results of Eqs. (8), (12), and (29). For this comparison the boundary and initial conditions are taken in the following way. The curve for Eq. (29) was drawn using the following data: $\rho = 480$ lbs./cu. ft., $H = -6.76$ BTU/lb., $k_2 = 25$ BTU/(hr.)(sq. ft.)(°F/ft.), $k_1 = 26$ BTU./(hr.)(sq. ft.)(°F/ft.), $c_{p_1} = .12$ BTU/(lb.)(°F), $c_{p_2} = .14$ BTU/(lb.)(°F). $g(t)$ was chosen to be $-117\, e^{-t}$ BTU/(sq. ft.)(hr.) and $\phi(x)$ to be $-50 \tan^{-1} x$ °F. This choice of $g(t)$ is such that the heat added per unit time decreases with increasing time, while $\phi(x)$ has the property that the initial temperature at $x = 0$ is the critical temperature and the temperature decreases with increasing x until it reaches 25π°F below critical at $x = \infty$. For the curve of Eq. (12) the data is acquired from that for Eq. (29) in the same manner that the solution (12) may be obtained from the solution (29), i.e., $k_2 = k$, $c_{p_2} = c_p$, $\phi(x) =$

$\phi(0) = 0$. The data for the curve of equation (8) is deduced from that for Eq. (12) similarly, i.e., $g(t) = g(0) = -117$ BTU/(sq. ft.)(hr.). Since only two terms of the solution (29) have been determined, all curves are drawn using only the first two terms of each series. The straight line shown in Figure 2 represents the time it would take a

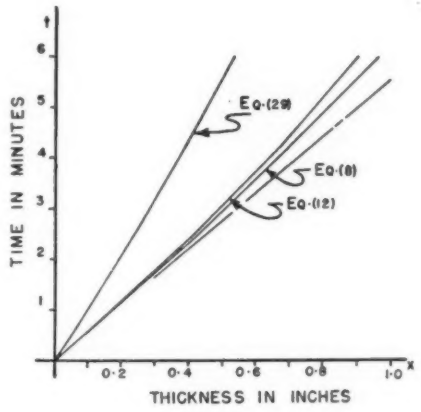

FIG. 2.

slab of thickness x at the critical temperature to recrystallize uniformly with a constant heat input of -117 BTU/(sq. ft.)(hr.).

According to (29), if t can be taken so small that all terms above the first are neglected, recrystallization cannot take place if $k_1\phi'(0) > k_2 g(0)$. In fact, this condition would represent decrystallization and the curve would be in the negative x half plane.

BOOK REVIEWS

Electromagnetic theory. Proceedings of Symposia in Applied Mathematics, Volume II. American Mathematical Society, New York, 1950. iii + 91 pp. $3.00.

This volume contains the proceedings of the Second Symposium in Applied Mathematics, held at the Massachusetts Institute of Technology in July 1948. Of the seventeen papers included, ten are presented in the volume by abstract.

The subject matter covers a wide range under the general topic of Electromagnetic Theory including related quantum theory and statistics. The papers in full are: H. Feshbach, *The new quantum electrodynamics.* J. L. Synge, *Electromagnetism without metric.* W. H. Watson, *Discontinuity in electromagnetism.* L. Infeld, *The factorization method and its application to differential equations in theoretical physics.* R. J. Duffin, *Nonlinear electrical networks.* C. L. Pekeris, *Ray theory vs. normal mode theory in wave propagation problems.* A. E. Heins, *Systems of Wiener-Hopf integral equations and their application to some boundary value problems in electromagnetic theory.*

W. D. HAYES

Deformation and flow. An elementary introduction to theoretical rheology. By Markus Reiner. H. K. Lewis & Co., Ltd., London, 1949. xix + 342 pp. 32s. 6d. (Sold in the U.S.A. by Interscience Publishers, New York, at $6.50).

This treatment of the theory of Rheology is primarily concerned with the fundamental behavior of the various rheological media and not with the general formulations of the typical problems of the mechanics of continua. Thus, the kinematic considerations have been limited to deformations which are essentially pure shear but the range of media discussed is very broad. There are chapters dealing with viscous fluids, plastic media, the generalized Newtonian fluid, and complex media such as concrete; and there are chapters devoted to discussions of work hardening, creep, strength and failure, and the behavior near solid boundaries. The book contains a wealth of information for those interested in the behavior of various media under comparatively simple loading or simple flow conditions.

G. F. Carrier

Experimental designs. By William G. Cochran and Gertrude M. Cox. John Wiley & Sons, Inc., New York and Chapman & Hall, Ltd., London. ix + 454 pp. $5.75.

In this book the authors present a large number of different designs covering a great variety of experimental situations. The application of these designs and their analysis is illustrated by well chosen examples. The book being intended as an elementary text and as a guide for research workers with little or no mathematical training, the authors do not fully discuss the basic assumptions underlying the analysis and the principles of statistical inference that lead from the assumptions to the analysis. However, the book distinguishes itself from many other elementary books in Statistics by giving a correct, although incomplete discussion of the necessary fundamental statistical principles.

The greatest value of the book in the reviewer's opinion is the extensive collection of experimental designs presented. A disadvantage for the reader with more mathematical background is the absence of closed expressions for the computation of test functions. Such closed expressions, which would inform the mathematically trained reader at a glance, are replaced by numerical examples, probably because formulae involving multiple summation signs and other formidable mathematical symbols would deter the average reader. The reviewer believes, however, that the book will be very profitable reading also for the student of Mathematical Statistics, particularly when used as complementary reading to a treatise presenting the mathematical theory of the design of experiments and the analysis of variance.

Henry B. Mann

An introduction to the LaPlace transformation. By J. C. Jaeger. John Wiley & Sons, Inc., New York and Methuen & Co., Ltd., London. viii + 132 pp. $1.50.

This member of the Methuen series of "Monographs on physical subjects" is concerned with the use of the LaPlace transform in solving differential equations with constant coefficients. The first chapter introduces the theorems needed; the second deals with the analysis of problems associated with electric circuit analysis; the third introduces a few more theorems and the final chapter applies the operational technique to a partial differential equation. The concise presentation seems well suited to those who wish to become familiar with an operational tool without wading through a careful analysis of the pertinent mathematics.

G. F. Carrier

Elements of aerodynamics of supersonic flows. By Antonio Ferri. The Macmillan Company, New York, 1949. x + 434 pp. $10.00.

This is a text designed for the practicing supersonic aerodynamicist. By limiting his field of coverage to purely supersonic flow, excepting incidentally subsonic flow where encountered behind a shock, the author has produced a treatment which may be termed reasonably comprehensive. The mathematical requirements for the reader are not too high—he needs a thorough knowledge of basic analysis through differential equations, but partial differential equations theory is not prerequisite and complex variables do not appear in the book. Although linearized theory is presented, the principal emphasis is on the nonlinear, as befits a realistic approach. Test results are included wherever feasible for comparison and illustration and the entire presentation is made very graphic by a satisfying number of figures. The index is completely inadequate.

This is definitely not a text for workers in fundamental research. The main reason for this is that the main fundamental problems of high speed flow lie outside the field of coverage of the book; they involve subsonic flow regions, viscous effects, and stability questions, and much of the mathematics involved is more advanced. Also, in designing his book for the engineer, the author has emphasized the specific at the expense of the fundamental, omitting such basic concepts in supersonic flow as those of the Euler momentum integral and the linearized and hypersonic similitudes.

W. D. Hayes

Methoden der praktischen analysis. By F. A. Willers. 2nd ed. Walter de Gruyter & Co., Berlin, 1950. 410 pp. DM 24.-.

The first edition of this well-known text appeared in 1928. In preparing the present revised edition, the author has included many new methods developed since the publication of the first edition. To keep the size of the volume approximately the same, some special graphical methods discussed were omitted. While there are frequent minor revisions throughout the book, the major additions are sections on the approximate integration of integral equations of the second kind, evaluation of improper integrals, trigonometric interpolation, and the solution of systems of linear equations by iteration. The last chapter on approximate integration of differential equations has been completely rewritten. According to a statement in the preface, the manuscript of this revised edition was completed in 1943. This explains the fact that useful techniques developed during the war in the United States and Great Britain have not been included.

W. Prager

Tables of the generalized exponential-integral functions. By The Staff of the Computation Laboratory. Harvard University Press, Cambridge, Massachusetts, 1949. xxv + 416 pp. $8.00.

The twenty-five introductory pages of this volume are concerned with the definitions of the functions to be tabulated, a discussion of the computation methods, the interpolation technique and a brief application discussion with bibliography. Three functions are computed: They are $\int_0^x F_n(a, s)\, ds$ where the three integrands used are $u^{-1}(1 - e^{-u})$, $u^{-1}e^{-u} \sin u$, and $u^{-1}(1 - e^{-u} \cos u)$, with $u(a, s) = (a^2 + s^2)^{1/2}$.

These functions are tabulated in the range $0 \leq a \leq .049$ in steps $\Delta a = .001$ with $0 \leq x \leq .049$ in steps of .001; in $0 \leq a \leq .098$ and $0 \leq x \leq .098$ with increments in x and a of .002; in $0 \leq a \leq .245$ and $0 \leq x \leq .245$ with increments of .005; for $0 \leq a \leq .49$, $0 \leq x \leq .49$ with increments .01; in $0 \leq a, x \leq .98$ with increments .02; in $0 \leq a, x \leq 2.45$ with increments .05; for $0 \leq a, x \leq .49$ with increments .1; and for $0 \leq a, x \leq .98$ with increments .2.

G. F. Carrier

Tables of the function $\sin \phi/\phi$ *and of its first eleven derivatives.* By The Staff of the Computation Laboratory. Harvard University Press, Cambridge, Massachusetts, 1949. sviii + 241 pp. $8.00.

The function $\sin \phi/\phi$ and its first eleven derivatives are tabulated for values of ϕ in the range $0 \leq \phi < 20\pi$. The increments in ϕ are taken to be .5 degrees for computation convenience. Remarks about computational procedure and a brief discussion of the applications in which these functions are useful make up the introductory pages.

G. F. CARRIER

Tables of the Bessel functions of the first kind of orders sixty-four through seventy-eight. By The Staff of the Computation Laboratory. Harvard University Press, Cambridge, Massachusetts, 1949. 566 pp. $8.00.

The Bessel Functions $J_{64}(x)$, $J_{65}(x)$, $\cdots$ $J_{78}(x)$ are tabulated to ten decimal places. The argument varies in steps of .01 for $0 \leq x < 100$.

G. F. CARRIER

The variational principles of mechanics. By Cornelius Lanczos. University of Toronto Press, Toronto, 1949. xxv + 307 pp. $5.75.

This is an excellent book covering one of the most brilliant achievements of scientific theory, and the text will certainly be welcomed by teachers and students of analytical mechanics.

The author leads the reader through the historical development, starting from the basic concepts of mechanics and discussing a number of specific examples to familiarize the student with the general principles. Short, clear summaries present the fundamental results obtained in each section and make it possible to follow step by step the general line of thought. The chapter on Liouville's theorem may be taken as a typical illustration of the author's method, with a general discussion of phase-space, then a few simple examples, and finally the general theorem about integral-invariants.

One may only regret that the author did not include Ehrenfest's adiabatic invariants, which play such an important role in quantum theories. The analytical methods used in the old quantum mechanics are very completely and clearly discussed in the last chapter, which will be found especially valuable by many students in theoretical physics.

This is a brilliant presentation of a great chapter of science, and the text will be of great help to students who have studied mechanics in Synge and Griffith's excellent treatise where these problems were barely sketched.

L. BRILLOUIN

The meaning of relativity. By Albert Einstein. Princeton University Press, Princeton, New Jersey, 1950. 150 pp. $2.50.

The first part of the book is a general explanation of the physical background of Relativity. It begins with a discussion of the fundamental notions of space and time in classical physics and the role of relativity in classical mechanics, where all physical laws are assumed to be invariant for Cartesian systems of space coordinates, while the time variable is not supposed to be affected by these Cartesian changes of coordinates; at this stage, time and distance are absolute.

This set of definitions fails when applied to the Maxwell-Lorentz equations of electromagnetism. The discussion of this problem led Einstein to the discovery of the theory of Special Relativity, and the trend of thought is very clearly explained in the second chapter of the book, where the fusion of space and time in a comprehensive system of 4 variables is introduced.

The famous relation $E = mc^2$ is obtained, but there is just a short footnote to tell of the experimental proofs of this most important prediction.

General relativity was developed in order to include gravitation in this general description. The point of departure is the identity of inert mass and gravitational mass. This leads directly to the principle of equivalence, which states that gravitation plays a similar role to acceleration, and explains the need for Riemannian geometry. The connection between the General Relativity and Mach's ideas makes for a very interesting discussion of the fundamental assumptions and of the new results to be expected. The chapter concludes in favor of a space-bounded universe, as representing the most satisfactory type of general equation.

In the appendix to the second edition, this cosmologic problem is discussed, in connection with the experimental results about the red shift of spectral lines from distant nebulae and the expanding universe. The author does not believe in the addition of a "cosmologic term" in the equations of gravity, and presents the theory of Friedman as the best solution, with a four-dimensional universe that is isotropic with respect to three dimensions (space). This yields a value of 10^9 years for the age of the universe, a very small value indeed and difficult to reconcile with other astronomical data.

The second appendix is a new addition to the present third edition, and contains the new generalized theory of gravitation, recently developed by A. Einstein. This is an attempt to include Electromagnetism in the general geometry of the universe, instead of leaving Maxwell's equations as a separate group of equations, unconnected with the geometry. The previous general Relativity was a frame in which Electromagnetism was included as a separate picture. The aim is now to rebuild frame and picture together as a single unit. For this problem, our general knowledge of physics does not permit an unequivocal choice as did the principle of equivalence for gravitation. The only clue is that something similar to Maxwell's equation must be contained within the total field.

Many such attempts have been presented in the last decades, and they can be classified into two main groups. The electromagnetic field representing a skew-symmetrical tensor, one may try to add such skew-symmetrical terms to either

A—the $g_{\mu\nu}$ fundamental metric tensor

or B—the $\Gamma^i_{\mu\nu}$ coefficients of parallel displacement

(the $\Gamma^i_{\mu\nu}$ do not constitute a tensor)

Both quantities were usually defined as symmetrical in $\mu\nu$, and can be completed with skew-symmetrical terms. Einstein uses both generalizations simultaneously and builds up the corresponding Riemann Tensor for curvature. He then looks for field equations, and in order to be certain of the compatibility of these equations he derives them from a variational principle according to the well-known Hamilton procedure. After some simplification, the result is a system of equations very similar to Maxwell's equations (with no free magnetic charges).

The author does not give any discussion of the practical value of these new equations and does not tell the new facts that they may include. One must remember that the theoretical problem was a formidable one, and that A. Einstein has been working on it for two decades, trying all possible lines of attack and attempting to use all sorts of new geometrical structures. He finally makes a choice, and gives us a set of formulas that seem to him to represent the best logical solution, leaving to others the trouble to discuss the practical and physical meaning of the new theory. Many questions remain still obscure, and first of all: how is it possible to obtain the equations of motion for different particles with different masses, charges and magnetic moments? One single geodesic can not do, since all trajectories are different from each other. This problem is not discussed in Einstein's paper, and we hope he shall give us his ideas about the type of logical solution to be expected.

The reader will refer with great interest to the paper published by Einstein in the "Scientific American", p. 13, vol. 182, no. 4, April 1950, where the philosophical background of the theory is discussed with great care and remarkable clarity. Einstein indicates in this paper a very curious feature of his theory: the initial conditions cannot be freely chosen. He points out that this fact does not seem to lead to serious trouble, provided one may prove that the system is extensive enough to cover all the possible physical problems, and this is still an open question.

Let us admire the genius of a great thinker and an exceptional scientist, who fulfills now, in his ripe age, an audacious plan of research that he started as a young man. A wonderful achievement and a rare one.

L. BRILLOUIN

Physical principles of oil production. By Morris Muskat. First Edition. McGraw-Hill Book Company, Inc., New York, Toronto, London, 1949. xv + 922 pp. $15.00.

As stated in the preface this work is neither a textbook nor a reorganized version of well established theories for pedagogical purposes, but an exposition of material which is in an active state of flux, with the purpose of formulating and correlating the physical principles and facts underlying the mechanisms of oil production.

The physics of oil production has been a rapidly expanding branch of applied physics in the last fifteen years and one cannot help but admire the art and ingenuity with which the scientific approach has been introduced in a field where rule of thumb used to prevail and where the object of study, the oil reservoir, offers an infinite variety of individual aspects which are susceptible to only limited observation.

Chapters 1 to 3 describe and discuss the general physical properties of petroleum fluids and oil bearing rocks with particular reference to the thermodynamic behavior of multiphase hydrocarbon systems and their interaction with porous rock. The following three chapters deal with the motion of single phase fluids through porous rock. Incompressible fluid flows form the most important class for practical applications and such flows in homogeneous and heterogeneous media are treated by solving the Laplace equation. Many types of flow fields are described corresponding to various physical conditions of stratification and well penetration. The effect of compressibility is also introduced and the corresponding diffusion-type equation is applied to a few idealized cases. The more difficult problems of mult phase fluid flow phenomena are treated in Chapters 7 to 11. The physical concepts and characteristics of this type of flow are reviewed and applied to the solution of several simple steady-state flow equations. This is supplemented by examples of actual performance records of fields. The last three chapters are devoted to problems of secondary recovery, condensate reservoirs, well spacing and the estimation of recoverable reserves. The mathematical and analogue computer methods developed for the solution of these problems will be of interest also to those who are not directly concerned with oil production.

In view of the authoritative nature of this work it is unfortunate that serious criticism may be expressed concerning the presentation of the general laws of fluid flow through porous media. The treatment lacks the clarity and generality which one expects to find in a treatise of this scope. In the opinion of the reviewer the book would have gained appreciably if it had followed a more systematic presentation as developed, e.g., by M. K. Hubbert (cf. Theory of Ground Water Motion, The Journal of Geology Vol. 48, No. 8, November-December 1940).

A great deal of interest has been attracted lately to the question of the rapid depletion of the world oil reserves. The challenge to the applied mathematician and physicist is well illustrated by this most thoroughly documented treatment of the field.

M. A. BIOT

3 . . . STATISTICAL INDEPENDENCE

by [illegible], *Columbia University*. One of Fréchet's leading studies presents the first [illegible] a new statistical theory which [illegible] and [illegible] as part of the [illegible] and [illegible] experimentation as part of the [illegible] (in certain [illegible] experiments [illegible] to be carried out in a single step); and it includes the general [illegible] from [illegible] previous studies, the methods [illegible] and to point an interval estimation. One of the WILEY PUBLICATIONS IN STATISTICS, Walter A. Shewhart, Editor. *August 1950. 170 pages. $5.00.*

3 . . . MATHEMATICS of RELATIVITY

by O. Y. Mayer, *University of Michigan*. Offers four distinct books in two volumes. The author concentrates [illegible] gravitation and electromagnetism [illegible] the original theory of curved space. The [illegible] author uses a distinctive approach which makes the presentation easy to follow without sacrificing rigor. He begins with the inverse square law of attraction and then introduces the difficult points (such as changes in fundamental concepts and refinements of mathematical technique) only after the need for them is made clear. A volume in the APPLIED MATHEMATICS SERIES, Edited by I. S. Sokolnikoff. *September 1950. 172 pages. $3.50.*

3 . . . RESPONSE of PHYSICAL SYSTEMS

by John D. Trimmer, *University of Tennessee*. Offers an approach to the problem of defining instrumentation. It is written from a point of view based on a certain pattern of experience. The point of view is: A system, of nature unspecified, is subject to an input, or "forcing," and gives an output, or "response." A general study, closely linked with cybernetics, is built up around this pattern. It includes not only physical systems, such as instruments, regulators, and servos, but also biological and sociological entities. This general study is called "system response." *August 1950. 268 pages. $5.00.*

Two New Carus Monographs

The THEORY of ALGEBRAIC NUMBERS

by Harry Pollard, *Cornell University*. No. 9 of the CARUS MONOGRAPHS. *1950. 143 pages. $3.00.*

The ARITHMETIC THEORY of QUADRATIC FORMS

By Burton W. Jones, *University of Colorado*. No. 10 of the CARUS MONOGRAPHS. *1950. 212 pages. $3.00.*

Send for copies on approval

JOHN WILEY & SONS, Inc.

440 Fourth Avenue — New York 16, N. Y.

THEORY OF FLOW AND FRACTURE OF SOLIDS. Volume I

by A. E. Green, Consulting Engineer, Westinghouse Research Laboratories, 372 pages, $10.00

Emphasis on the mathematical methods and mechanical laws governing the perfection and the frontiers of knowledge of solids, designed to acquaint the student and workers in engineering and the allied arts with the theory and results of investigations to important practical problems.

INTERNAL BALLISTICS OF ROCKETS

by R. N. Wimpress, Industrial Engineering, Inc., Los Angeles, California. 211 pages, $4.50

A general treatment of the experimental data presented in this book is primarily concerned with the utilization of solid propellant double-base smokeless powder in rockets of relatively short burning time, although there is some information on rockets of longer burning time (up to one minute) and on other types of solid propellants. Ignition, nozzle design, heating of the motor walls, and instrumentation are also covered.

ENGINEERING MECHANICS. New 3rd Edition

by S. Timoshenko and D. H. Young, Stanford University. Ready in November

Includes approximately 300 new problems, the new edition of this well-known text is, as before, to build a strong foundation in this fundamental subject, to acquaint the student with the many general methods of attacking problems and to illustrate the application of these methods to practical engineering problems.

Send for copies on approval

D. VAN NOSTRAND-REIN HOLD BOOK COMPANY, Inc.

250 West 52nd Street New York 19, N. Y.